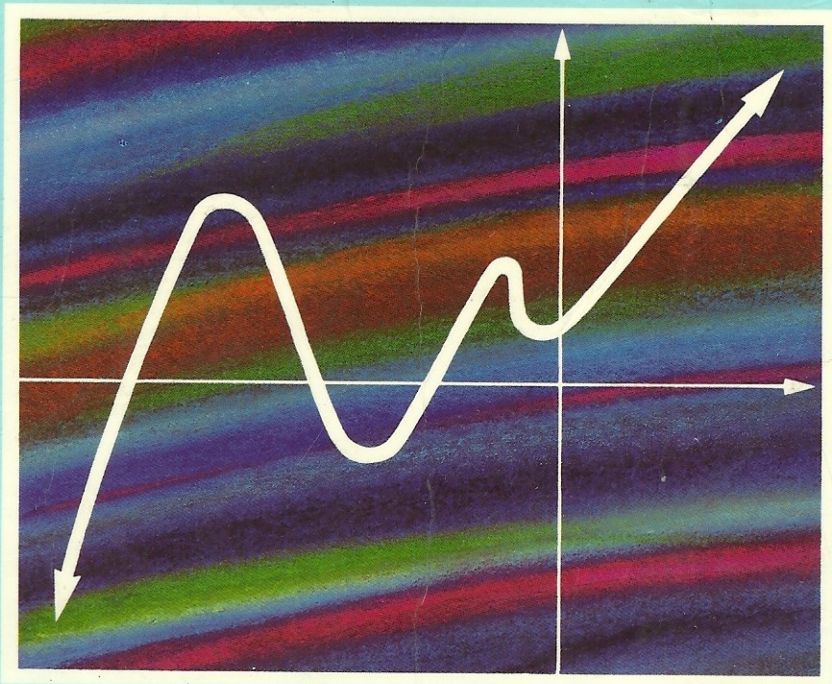
Schaum

ALGEBRA LINEAL

SEGUNDA EDICION

Seymour Lipschutz



McGraw Hill

ALGEBRA LINEAL

Segunda edición

SEYMOUR LIPSCHUTZ, Ph. D.
Temple University

Traducción:
CELIA MARTINEZ ONTALBA

Revisión:
LORENZO ABELLANAS
Catedrático Métodos Matemáticos de la Física
Universidad Complutense de Madrid

McGraw-Hill

MADRID • BUENOS AIRES • CARACAS • GUATEMALA • LISBOA • MEXICO
NUEVA YORK • PANAMA • SAN JUAN • SANTAFE DE BOGOTA • SANTIAGO • SAO PAULO
AUCKLAND • HAMBURGO • LONDRES • MILAN • MONTREAL • NUEVA DELHI • PARIS
SAN FRANCISCO • SIDNEY • SINGAPUR • ST. LOUIS • TOKIO • TORONTO

ALGEBRA LINEAL. Segunda edición

Edificio Oasis-A, 1.ª planta
Basauri, s/n
28023 Aravaca (Madrid)

Traducido de la segunda edición en inglés de
LINEAR ALGEBRA

ISBN: 0-07-038007-4

ISBN: 84-7615-758-4
Depósito legal: M. 119 / 1993

Cubierta: Juan García
Compuesto e impreso en: Fernández Ciudad, S. L.

PRINTED IN SPAIN - IMPRESO EN ESPAÑA

Contenido

Prólogo

El álgebra lineal se ha convertido en los últimos años en una parte esencial de los conocimientos matemáticos requeridos a matemáticos, ingenieros, físicos y otros científicos. Este requerimiento refleja la importancia y la amplitud de sus aplicaciones.

Este libro se ha proyectado como libro de texto en un curso regular de álgebra lineal o como suplemento a los textos clásicos en uso. Su propósito es presentar una introducción al álgebra lineal que todos los lectores encuentren provechosa, cualquiera que sea su campo de especialización. Se ha incluido más material del que puede abarcarse en muchos primeros cursos, con objeto de hacer el texto más flexible, proporcionar un libro de referencia útil y estimular el interés futuro en el tema.

Cada capítulo comienza con enunciados claros de las definiciones, principios y teoremas pertinentes, junto con ejemplos y otro material descriptivo. A esto siguen colecciones graduadas de problemas resueltos y problemas suplementarios. Los problemas resueltos sirven para ilustrar y ampliar la teoría, concretan aquellos puntos sutiles cuyo desconocimiento lleva al estudiante a sentirse continuamente en un terreno inseguro, y suministran la repetición de los principios básicos tan vital para un aprendizaje efectivo. Entre los problemas resueltos se incluyen numerosas demostraciones de teoremas. Los problemas suplementarios sirven como revisión completa del material de cada capítulo.

El primer capítulo trata los sistemas de ecuaciones lineales. Esto proporciona la motivación y las herramientas de cálculo básicas para el material subsiguiente. Tras haber introducido los vectores y las matrices, aparecen capítulos sobre espacios vectoriales y subespacios, y sobre productos internos. Siguen capítulos que cubren determinantes, valores propios y vectores propios, y diagonalización de matrices (bajo similaridad) y formas cuadráticas (bajo congruencia). Los capítulos posteriores abarcan las aplicaciones lineales abstractas y sus formas canónicas, específicamente las formas canónicas triangular, de Jordan y racional. El último capítulo trata las aplicaciones lineales abstractas en espacios con producto interno.

Los principales cambios en la segunda edición han sido por razones pedagógicas (de forma) más que de contenido. Aquí la noción de aplicación matricial se introduce al principio del texto,

y los productos internos justo después del capítulo sobre espacios vectoriales y subespacios. Asimismo, se presentan los algoritmos de reducción por filas, inversión de matrices, cálculo de determinantes y diagonalización de matrices y formas cuadráticas, usando notación algorítmica. Además, temas tales como matrices elementales, factorización LU, coeficientes de Fourier y varias normas en $\mathbf{R}^n$ se introducen directamente en el texto, en lugar de hacerlo en las secciones de problemas. Por último, al tratar los aspectos abstractos más avanzados en la parte final del texto, facilitamos el uso de éste en un curso elemental o en uno de dos semestres de álgebra lineal.

Deseo dar las gracias al personal de McGraw-Hill Schaum Series, especialmente a John Aliano, David Beckwith y Margaret Tobin, por sus inestimables sugerencias y su cooperación, verdaderamente útil. Por último, quiero expresar mi gratitud a Wilhelm Magnus, mi maestro, consejero y amigo, que me introdujo en la belleza de las matemáticas.

Temple University
Enero de 1991

SEYMOUR LIPSCHUTZ

CAPITULO 1

Sistemas de ecuaciones lineales

1.1. INTRODUCCION

La teoría de las ecuaciones lineales juega un papel importante y motivador en el ámbito del álgebra lineal. De hecho, muchos problemas de álgebra lineal son equivalentes al estudio de un sistema de ecuaciones lineales, como hallar el núcleo de una aplicación lineal o caracterizar el subespacio generado por un conjunto de vectores. Por ello, las técnicas introducidas en este capítulo serán aplicables al tratamiento más abstracto dado posteriormente. Por otra parte, algunos de los resultados del tratamiento abstracto arrojarán nueva luz sobre la estructura de los sistemas de ecuaciones lineales «concretos».

Este capítulo investiga sistemas de ecuaciones lineales y describe detalladamente el algoritmo de eliminación gaussiano, que se utiliza para hallar su solución. Aunque serán estudiadas en detalle en el Capítulo 3, las matrices, junto con ciertas operaciones entre ellas, se introducen también aquí, puesto que están estrechamente relacionadas con los sistemas de ecuaciones lineales y su solución.

Todas nuestras ecuaciones involucrarán números específicos denominados *constantes* o *escalares*. Para simplificar, en este capítulo asumimos que todos nuestros escalares pertenecen al cuerpo de los números reales **R**. Las soluciones de nuestras ecuaciones también involucrarán n-plas $u = (k_1, k_2, \ldots, k_n)$ de números reales llamadas *vectores*. El conjunto de tales n-plas se denota por $\mathbf{R}^n$.

Apuntamos que los resultados de este capítulo son válidos asimismo para ecuaciones sobre el cuerpo complejo **C**, o sobre cualquier cuerpo arbitrario K.

1.2. ECUACIONES LINEALES. SOLUCIONES

Por una *ecuación lineal* con incógnitas $x_1, x_2, \ldots, x_n$ entendemos una ecuación que puede escribirse en la *forma convencional*:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad [1.1]$$

donde $a_1, a_2, \ldots, a_n, b$ son constantes. La constante a_k se denomina el *coeficiente* de x_k y b se denomina la *constante* de la ecuación.

Una *solución* de la ecuación lineal anterior es un conjunto de valores de las incógnitas, digamos $x_1 = k_1, x_2 = k_2, \ldots, x_n = k_n$, o simplemente una n-pla $u = (k_1, k_2, \ldots, k_n)$ de constantes, con la propiedad de que es cierta la siguiente expresión (obtenida sustituyendo cada x_i por k_i en la ecuación):

$$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$$

Se dice entonces que este conjunto de valores *satisface* la ecuación.

El conjunto de todas las soluciones se llama *conjunto solución*, *solución general* o, simplemente, la *solución* de la ecuación.

Nota: Las nociones anteriores presuponen implícitamente que existe un orden entre las incógnitas. Con el fin de evitar los subíndices, normalmente utilizaremos variables x, y, z para denotar tres incógnitas, x, y, z, t para denotar cuatro incógnitas y x, y, z, s, t para denotar cinco incógnitas, considerándolas ordenadas.

EJEMPLO 1.1

a) La ecuación $2x-5y+3xz=4$ no es lineal, porque el producto de dos incógnitas es de segundo grado.

b) La ecuación $x + 2y - 4z + t = 3$ es lineal en las cuatro incógnitas x, y, x, t.

La 4-pla $u = (3, 2, 1, 0)$ es una solución de la ecuación porque

$$3 + 2(2) - 4(1) + 0 = 3 \qquad \text{o} \qquad 3 = 3$$

es una expresión cierta. Sin embargo, la 4-pla $v = (1, 2, 4, 5)$ no es una solución de la ecuación porque

$$1 + 2(2) - 4(4) + 5 = 3 \qquad \text{o} \qquad -6 = 3$$

no es cierto.

ECUACIONES LINEALES CON UNA INCOGNITA

El siguiente resultado básico está probado en el Problema 1.5.

Teorema 1.1: Consideremos la ecuación lineal $ax = b$.

i) Si $a \neq 0$, $x = b/a$ es solución única de $ax = b$.

ii) Si $a = 0$, pero $b \neq 0$, $ax = b$ no tiene solución.

iii) Si $a = 0$ y $b = 0$, todo escalar k es solución de $ax = b$.

EJEMPLO 1.2

a) Resolvamos $4x - 1 = x + 6$.

Trasponemos para obtener la ecuación en forma convencional: $4x - x = 6 + 1$ ó $3x = 7$. Multiplicamos por $1/3$ para obtener la solución única $x = \frac{7}{3}$ [Teorema 1.1 i)].

b) Resolvamos $2x - 5 - x = x + 3$.
Reescribimos la ecuación en forma convencional: $x - 5 = x + 3$, o $x - x = 3 + 8$, o $0x = 8$. La ecuación no tiene solución [Teorema 1.1 ii)].

c) Resolvamos $4 + x - 3 = 2x + 1 - x$.
Reescribimos la ecuación en forma convencional: $x + 1 = x + 1$, o $x - x = 1 - 1$, o $0x = 0$. Todo escalar k es una solución [Teorema 1.1 iii)].

ECUACIONES LINEALES DEGENERADAS

Una ecuación lineal se dice *degenerada* si tiene la forma

$$0x_1 + 0x_2 + \cdots + 0x_n = b$$

esto es, si cada coeficiente es igual a cero. La solución de tal ecuación se halla como sigue:

Teorema 1.2: Consideremos la ecuación lineal degenerada $0x_1 + 0x_2 + \cdots + 0x_n = b$.

i) Si $b \neq 0$, la ecuación no tiene solución.
ii) Si $b = 0$, todo vector $u = (k_1, k_2, \ldots, k_n)$ es una solución.

Demostración. i) Sea $u = (k_1, k_2, \ldots, k_n)$ un vector cualquiera. Supongamos $b \neq 0$. Sustituyendo u en la ecuación obtenemos:

$$0k_1 + 0k_2 + \cdots + 0k_n = b \qquad \text{o} \qquad 0 + 0 + \cdots + 0 = b \qquad \text{o} \qquad 0 = b$$

No es una expresión cierta porque $b \neq 0$. Por tanto, ningún vector u es solución.

ii) Supongamos $b = 0$. Sustituyendo u en la ecuación obtenemos:

$$0k_1 + 0k_2 + \cdots + 0k_n = 0 \qquad \text{o} \qquad 0 + 0 + \cdots + 0 = 0 \qquad \text{o} \qquad 0 = 0$$

que es una expresión cierta. Por consiguiente, todo vector u en $\mathbf{R}^n$ es una solución, como se pretendía.

EJEMPLO 1.3. Describir la solución de $4y - x - 3y + 3 = 2 + x - 2x + y + 1$.
La reescribimos en forma convencional agrupando términos y trasponiendo:

$$y - x + 3 = y - x + 3 \qquad \text{o} \qquad y - x - y + x = 3 - 3 \qquad \text{o} \qquad 0x + 0y = 0$$

La ecuación es degenerada con constante nula; por tanto, todo vector $u = (a, b)$ en $\mathbf{R}^2$ es una solución.

ECUACIONES LINEALES NO DEGENERADAS. PRIMERA INCOGNITA

Esta subsección trata la solución de una sola ecuación lineal no degenerada con una o más incógnitas, digamos

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

Por la *primera incógnita* en tal ecuación entendemos la primera con coeficiente no nulo. Su posición p en la ecuación es entonces el menor valor entero de j para el cual $a_j \neq 0$. En otras palabras, x_p es la primera incógnita si $a_j = 0$ para $j < p$, pero $a_p \neq 0$.

EJEMPLO 1.4. Consideremos la ecuación lineal $5y - 2z = 3$. Aquí y es la primera incógnita. Si las incógnitas son x, y y z, entonces $p = 2$ es su posición, pero si y y z son las únicas incógnitas, es $p = 1$.

El siguiente teorema, probado en el Problema 1.9, es aplicable.

Teorema 1.3: Consideremos una ecuación lineal no degenerada $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ con primera incógnita x_p.

i) Cualquier conjunto de valores de las incógnitas x_j con $j \neq p$ dará una única solución de la ecuación. (Las incógnitas x_j se llaman *variables libres* porque se les puede asignar cualquier valor.)
ii) Toda solución de la ecuación se obtiene en i).

(El conjunto de todas las soluciones se llama *solución general* de la ecuación.)

EJEMPLO 1.5

a) Hallar tres soluciones particulares de la ecuación $2x - 4y + z = 8$.

Aquí x es la primera incógnita. De acuerdo con ello, asignamos valores cualesquiera a las variables libres y y z y entonces despejamos x para obtener una solución. Por ejemplo:

1. Tomemos $y = 1$ y $z = 1$. La sustitución en la ecuación proporciona

$$2x - 4(1) + 1 = 0 \quad \text{o} \quad 2x - 4 + 1 = 0 \quad \text{o} \quad 2x = 11 \quad \text{o} \quad x = \tfrac{11}{2}$$

Entonces $u_1 = (\frac{11}{2}, 1, 1)$ es una solución.

2. Tomemos $y = 1$, $z = 0$. La sustitución proporciona $x = 6$. Por consiguiente, $u_2 = (6, 1, 0)$ es una solución.
3. Tomemos $y = 0$, $z = 1$. La sustitución proporciona $x = \frac{7}{2}$. Por tanto, $u_3 = (\frac{7}{2}, 0, 1)$ es una solución.

b) La solución general de la ecuación anterior, $2x - 4y + z = 8$, se obtiene como sigue:

En primer lugar, asignamos valores arbitrarios (llamados *parámetros*) a las variables libres, digamos $y = a$ y $z = b$. A continuación sustituimos en la ecuación para obtener

$$2x - 4a + b = 8 \quad \text{o} \quad 2x = 8 + 4a - b \quad \text{o} \quad x = 4 + 2a - \tfrac{1}{2}b$$

Entonces

$$x = 4 + 2a - \tfrac{1}{2}b, \; y = a, \; z = b \quad \text{o} \quad u = (4 + 2a - \tfrac{1}{2}b, a, b)$$

es la solución general.

1.3. ECUACIONES LINEALES CON DOS INCOGNITAS

Esta sección considera el caso especial de las ecuaciones lineales con dos incógnitas, x e y, esto es, ecuaciones que se pueden escribir en la forma convencional

$$ax + by = c$$

donde a, b y c son números reales. (Supondremos también que la ecuación es no degenerada, esto es, que a y b no son ambos nulos.) Cada solución de la ecuación es un par de números reales, $u = (k_1, k_2)$, que puede hallarse asignando un valor arbitrario a x y despejando y, o viceversa.

Toda solución $u = (k_1, k_2)$ de la ecuación anterior determina un punto en el plano cartesiano $\mathbf{R}^2$. Como a y b no son ambos nulos, todas las soluciones tales corresponden precisamente a los puntos de una línea recta (de ahí el nombre de «ecuación lineal»). Esta línea se denomina el *gráfico* de la ecuación.

EJEMPLO 1.6. Consideremos la ecuación lineal $2x + y = 4$. Encontramos tres soluciones de la ecuación como sigue. Primero escogemos un valor arbitrario para cualquiera de las incógnitas, digamos $x = -2$. Sustituimos $x = -2$ en la ecuación y obtenemos

$$2(-2) + y = 4 \qquad \text{o} \qquad -4 + y = 4 \qquad \text{o} \qquad y = 8$$

Entonces $x = -2$, $y = 8$, o sea, el punto $(-2, 8)$ en $\mathbf{R}^2$ es una solución. Ahora hallamos el corte con el eje y, esto es, sustituimos $x = 0$ en la ecuación para obtener $y = 4$. Por consiguiente, el punto $(0, 4)$ en el eje y es una solución. A continuación encontramos el corte con el eje x, esto es, sustituimos $y = 0$ en la ecuación para obtener $x = 2$. Por tanto, $(2, 0)$ en el eje x es una solución.

Para dibujar el gráfico de la ecuación, primero dibujamos las tres soluciones, $(-2, 8)$, $(0, 4)$ y $(2, 0)$, en el plano $\mathbf{R}^2$ como se muestra en la Figura 1-1. Después trazamos la línea L determinada por dos de las soluciones y constatamos que la tercera yace en L también. (De hecho, L es el conjunto de todas las soluciones de la ecuación.) La línea L es el gráfico de la ecuación.

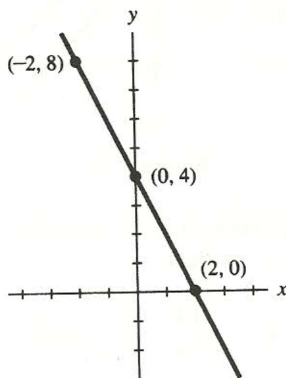


Gráfico de $2x+y=4$

Figura 1-1.

SISTEMA DE DOS ECUACIONES CON DOS INCOGNITAS

Esta subsección considera un sistema de dos ecuaciones lineales (no degeneradas) con las dos incógnitas x e y:

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned} \qquad [1.2]$$

(Por tanto, a_1 y b_1 no son simultáneamente nulos, ni tampoco lo son a_2 y b_2.) Este sistema simple se trata por separado porque tiene una interpretación geométrica, y porque sus propiedades motivan el caso general.

Un par de números reales $u = (k_1, k_2)$ que satisface ambas ecuaciones se llama una solución simultánea de las ecuaciones dadas, o una solución del sistema de ecuaciones. Existen tres casos, que pueden describirse geométricamente.

1. El sistema tiene exactamente una solución. Aquí los gráficos de las ecuaciones lineales se cortan en un punto, como en la Figura 1-2(*a*).
2. El sistema no tiene soluciones. Aquí los gráficos de las ecuaciones lineales son paralelos, como en la Figura 1-2(*b*).
3. El sistema tiene un número infinito de soluciones. Aquí los gráficos de las ecuaciones lineales coinciden, como en la Figura 1-2(*c*).

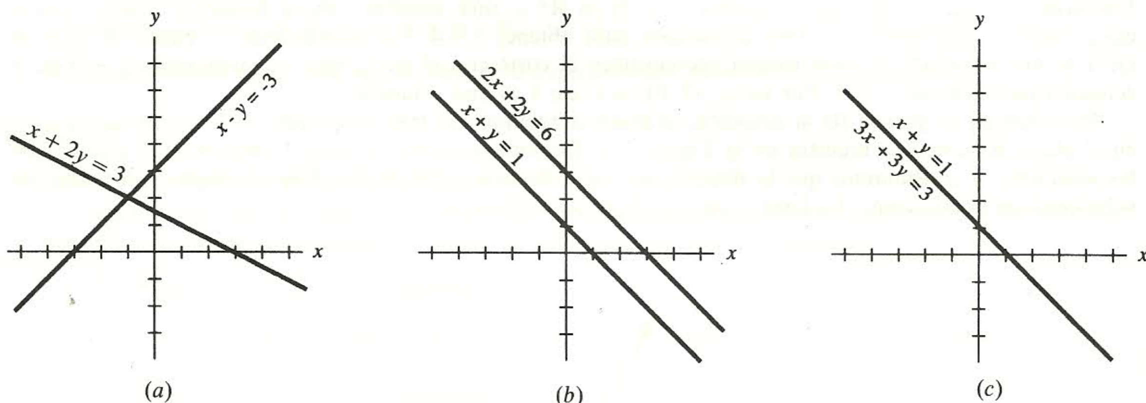


Figura 1-2.

Los casos especiales 2 y 3 sólo pueden ocurrir cuando los coeficientes de x e y en las dos ecuaciones lineales son proporcionales, es decir,

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} \qquad \text{o} \qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 = 0$$

En concreto, el caso 2 o el 3 ocurre si

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2} \qquad \text{o} \qquad \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$$

respectivamente. A menos que se establezca o sobrentienda otra cosa, suponemos que se trata con el caso general 1.

Nota: La expresión $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, que vale $a_1 b_2 - a_2 b_1$, se denomina *determinante* de orden dos. Los determinantes se estudiarán en el Capítulo 7. El sistema tiene una solución única cuando el determinante de los coeficientes es no nulo.

ALGORITMO DE ELIMINACION

La solución del sistema [1.2] puede obtenerse mediante el proceso conocido como eliminación, por medio del cual reducimos el sistema a una ecuación sencilla con sólo una incógnita. Aceptando que el sistema tiene solución única, este algoritmo de eliminación consiste en los dos pasos siguientes:

Paso 1. Sumar un múltiplo de una de las ecuaciones a la otra (o a un múltiplo no nulo de la otra) de forma que una de las incógnitas se elimina en la nueva ecuación.

Paso 2. Resolver la nueva ecuación para la incógnita dada y sustituir su valor en una de las ecuaciones originales para obtener el valor de la otra incógnita.

EJEMPLO 1.7

a) Consideremos el sistema

$$\begin{aligned} L_1: &\quad 2x + 5y = 8 \\ L_2: &\quad 3x - 2y = -7 \end{aligned}$$

Eliminamos x construyendo la nueva ecuación $L = 3L_1 - 2L_2$; esto es, multiplicando L_1 por 3 y L_2 por -2 y sumando las ecuaciones resultantes:

$$\begin{array}{ll} 3L_1: & 6x + 15y = 24 \\ -2L_2: & -6x + 4y = 14 \\ \hline \text{Suma:} & 19y = 38 \end{array}$$

Resolviendo la nueva ecuación para y se obtiene $y = 2$. Sustituyendo $y = 2$ en una de las ecuaciones originales, digamos en L_1, se obtiene

$$2x + 5(2) = 0 \quad \text{o} \quad 2x + 10 = 8 \quad \text{o} \quad 2x = -2 \quad \text{o} \quad x = -1$$

Entonces $x = -1$ e $y = 2$, o sea, el par $(-1, 2)$ es la solución única del sistema.

b) Consideremos el sistema

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = 5 \end{aligned}$$

Eliminamos x de las ecuaciones multiplicando L_1 por 2 y sumándola a L_2; es decir, formando la ecuación $L = 2L_1 + L_2$. Esto nos conduce a la nueva ecuación $0x + 0y = 13$. Es una ecuación degenerada, con constante no nula; en consecuencia, el sistema no tiene solución. (Geométricamente hablando, las líneas son paralelas.)

c) Consideremos el sistema

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = -8 \end{aligned}$$

Eliminamos x multiplicando L_1 por 2 y sumándola a L_2. Esto nos proporciona la nueva ecuación $0x + 0y = 0$, que es una ecuación degenerada en la que el término constante es nulo también. Por tanto, el sistema tiene un número infinito de soluciones, que corresponden a soluciones de cada ecuación. (Geométricamente hablando, las líneas coinciden.) Para encontrar la solución general

hacemos $y = a$ y sustituimos en L_1 obteniendo $x - 3a = 4$ o $x = 3a + 4$. De acuerdo con esto, la solución general del sistema es

$$(3a + 4, a)$$

donde a es un número real arbitrario.

1.4. SISTEMAS DE ECUACIONES LINEALES. SISTEMAS EQUIVALENTES. OPERACIONES ELEMENTALES

Esta sección considera un sistema de m ecuaciones lineales, digamos $L_1, L_2, \ldots, L_m$, con n incógnitas $x_1, x_2, \ldots, x_n$ que puede ponerse en la *forma convencional*

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad [1.3]$$

donde las a_{ij}, b_i son constantes.

Una *solución* (o *solución particular*) del sistema anterior es un conjunto de valores de las incógnitas, digamos $x_1 = k_1$, $x_2 = k_2$, ..., $x_n = k_n$, o una n-pla $u = (k_1, k_2, \ldots, k_n)$ de constantes, que es solución de cada una de las ecuaciones del sistema. El conjunto de todas las soluciones tales se denomina el *conjunto solución* o la *solución general* del sistema.

EJEMPLO 1.8. Considérese el sistema

$$\begin{array}{c} x_1 + 2x_2 - 5x_3 + 4x_4 = 3 \\ 2x_1 + 3x_2 + x_3 - 2x_4 = 1 \end{array}$$

Determinar si $x_1 = -8$, $x_2 = 4$, $x_3 = 1$, $x_4 = 2$ es una solución del sistema.

Sustituimos en cada ecuación para obtener

1. $-8 + 2(4) - 5(1) + 4(2) = 3$ o $-8 + 8 - 5 + 8 = 3$ o $3 = 3$
2. $2(-8) + 3(4) + 1 - 2(2) = 1$ o $-16 + 12 + 1 - 4 = 1$ o $-7 = 3$

No, no es una solución del sistema porque no es solución de la segunda ecuación.

SISTEMAS EQUIVALENTES. OPERACIONES ELEMENTALES

Se dice que dos sistemas de ecuaciones lineales con las mismas incógnitas son *equivalentes* si tienen el mismo conjunto solución. Una forma de producir un sistema que sea equivalente a uno dado, con ecuaciones lineales $L_1, L_2, \ldots, L_m$, es efectuar una sucesión de las siguientes operaciones, llamadas *operaciones elementales*:

[E_1] Intercambiar las ecuaciones i-ésima y j-ésima: $L_i \leftrightarrow L_j$.

[E_2] Multiplicar la ecuación i-ésima por un escalar no nulo k: $kL_i \to L_i$, $k \neq 0$.

[E_3] Sustituir la ecuación i-ésima por ella misma más k veces la j-ésima: $(kL_j + L_i) \to L_i$.

En la práctica efectuamos $[E_2]$ y $[E_3]$ en un solo paso, o sea, la operación

[E] Sustituir la ecuación i-ésima por k (no nulo) veces ella misma más k' veces la j-ésima:

$$(k'L_j + kL_i) \rightarrow L_i,\ k \neq 0$$

Lo anterior se enuncia formalmente en el siguiente teorema, probado en el Problema 1.46.

Teorema 1.4: Supongamos que un sistema de ecuaciones lineales (#) se obtiene de otro (*) mediante una sucesión finita de operaciones elementales. Entonces (#) y (*) tienen el mismo conjunto solución.

Nuestro método para resolver el sistema de ecuaciones lineales [1.3] consta de dos pasos:

Paso 1. Usar las operaciones elementales anteriores para reducir el sistema a uno equivalente más simple (en forma triangular o escalonada).

Paso 2. Usar la sustitución hacia atrás para hallar la solución del sistema más simple.

Los dos pasos se ilustran en el Ejemplo 1.9. En cualquier caso, por razones pedagógicas, discutimos en detalle primero el Paso 2 en la Sección 1.5 y luego el Paso 1 en la Sección 1.6.

EJEMPLO 1.9. La solución del sistema

$$\begin{aligned} x + 2y - 4z &= -4 \\ 5x + 11y - 21z &= -22 \\ 3x - 2y + 3z &= 11 \end{aligned}$$

se obtiene como sigue:

Paso 1. Primero eliminamos x de la segunda ecuación mediante la operación elemental $(-5L_1 + L_2) \rightarrow L_2$, esto es, multiplicando L_1 por -5 y sumándola a L_2; entonces eliminamos x de la tercera ecuación efectuando la operación elemental $(-3L_1 + L_3) \rightarrow L_3$, es decir, multiplicando L_1 por -3 y sumándola a L_3:

$$\begin{array}{rr} -5 \times L_1: & -5x - 10y + 20z = 20 \\ L_2: & 5x + 11y - 21z = -22 \\ \hline \text{nueva } L_2: & y - z = -2 \end{array} \qquad \begin{array}{rr} -3 \times L_1: & -3x - 6y + 12z = 12 \\ L_3: & 3x - 2y + 3z = 11 \\ \hline \text{nueva } L_3: & -8y + 15z = 23 \end{array}$$

Por tanto, el sistema original es equivalente al sistema

$$\begin{aligned} x + 2y - 4z &= -4 \\ y - z &= -2 \\ -8y + 15z &= 23 \end{aligned}$$

A continuación eliminamos y de la tercera ecuación aplicando $(8L_2 + L_3) \rightarrow L_3$, esto es, multiplicando L_2 por 8 y sumándola a L_3:

$$\begin{array}{rr} 8 \times L_2: & 8y - 8z = -16 \\ L_3: & -8y + 15z = 23 \\ \hline \text{nueva } L_3: & 7z = 7 \end{array}$$

Por consiguiente, obtenemos el siguiente sistema triangular equivalente:

$$\begin{aligned} x + 2y - 4z &= -4 \\ y - z &= -2 \\ 7z &= 7 \end{aligned}$$

Paso 2. Ahora resolvemos el sistema triangular más simple mediante sustitución hacia atrás. La tercera ecuación da $z = 1$. Sustituimos $z = 1$ en la segunda ecuación obteniendo

$$y - 1 = -2 \qquad \text{o} \qquad y = -1$$

Ahora sustituimos $z = 1$ e $y = -1$ en la primera ecuación para obtener

$$x + 2(-1) - 4(1) = -4 \qquad \text{o} \qquad x - 2 - 4 = -4 \qquad \text{o} \qquad x - 6 = -4 \qquad \text{o} \qquad x = 2$$

Entonces $x = 2$, $y = -1$, $z = 1$, o, en otras palabras, la terna ordenada $(2, -1, 1)$ es la solución única del sistema dado.

El anterior algoritmo de dos pasos para resolver un sistema de ecuaciones lineales se denomina *eliminación gaussiana.* El siguiente teorema se utilizará en el Paso 1 del algoritmo.

Teorema 1.5: Supongamos que un sistema de ecuaciones lineales contiene la ecuación degenerada

$$L: \quad 0x_1 + 0x_2 + \cdots + 0x_n = b$$

a) Si $b = 0$, L puede suprimirse del sistema sin alterar el conjunto solución.

b) Si $b \neq 0$, el sistema no tiene solución.

Demostración. Se obtiene directamente del Teorema 1.2. Esto es, la parte *a*) se obtiene del hecho de que todo vector en $\mathbf{R}^n$ es solución de L_1, y la parte *b*) del hecho de que L no tiene solución y por tanto tampoco la tiene el sistema.

1.5. SISTEMAS EN FORMA TRIANGULAR Y ESCALONADA

Esta sección considera dos tipos simples de sistemas de ecuaciones lineales: sistemas en forma triangular y el caso más general de sistemas en forma escalonada.

FORMA TRIANGULAR

Un sistema de ecuaciones lineales está en *forma triangular* si el número de ecuaciones es igual al número de incógnitas y si x_k es la primera incógnita de la k-ésima ecuación. Por tanto, un sistema de ecuaciones lineales triangular tiene la forma siguiente:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + \quad a_{1,n-1}x_{n-1} + \quad a_{1n}x_n &= b_1 \\ a_{22}x_2 + \cdots + \quad a_{2,n-1}x_{n-1} + \quad a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n &= b_{n-1} \\ a_{nn}x_n &= b_n \end{aligned} \qquad [1.4]$$

donde $a_{11} \neq 0$, $a_{22} \neq 0$, ..., $a_{nn} \neq 0$.

El sistema de ecuaciones lineales triangular anterior tiene una solución única, que puede obtenerse mediante el siguiente procedimiento, conocido como sustitución hacia atrás. Primero, resolvemos la última ecuación para la última incógnita, x_n:

$$x_n = \frac{b_n}{a_{nn}}$$

Segundo, sustituimos este valor de x_n en la penúltima ecuación y la resolvemos para la penúltima incógnita, x_{n-1}:

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}(b_n/a_{nn})}{a_{n-1,n-1}}$$

Tercero, sustituimos estos valores de x_n y x_{n-1} en la antepenúltima ecuación y la resolvemos para la antepenúltima incógnita, x_{n-2}:

$$x_{n-2} = \frac{b_{n-2} - (a_{n-2,n-1}/a_{n-1,n-1})[b_{n-1} - a_{n-1,n}(b_n/a_{nn})] - (a_{n-2,n}/a_{nn})b_n}{a_{n-2,n-2}}$$

En general, determinamos x_k sustituyendo los valores previamente obtenidos de x_n, x_{n-1}, ..., x_{k+1} en la ecuación k-ésima ecuación:

$$x_k = \frac{b_k - \sum_{m=k+1}^{n} a_{km} x_m}{a_{kk}}$$

El proceso finaliza cuando hemos determinado x_1. La solución es única puesto que, en cada paso del algoritmo, el valor de x_k está, por el Teorema 1.1 i), unívocamente determinado.

EJEMPLO 1.10. Consideremos el sistema

$$\begin{aligned} 2x + 4y - z &= 11 \\ 5y + z &= 2 \\ 3z &= -9 \end{aligned}$$

Como el sistema está en forma triangular, puede resolverse mediante sustitución hacia atrás.

i) La última ecuación proporciona $z = -3$.

ii) Sustituimos en la segunda ecuación para obtener $5y - 3 = 2$ ó $5y = 5$ ó $y = 1$.

iii) Sustituimos $z = -3$ e $y = 1$ en la primera ecuación para obtener

$$2x + 4(1) - (-3) = 11 \quad \text{o} \quad 2x + 4 + 3 = 11 \quad \text{o} \quad 2x = 4 \quad \text{o} \quad x = 2$$

Entonces el vector $u = (2, 1, -3)$ es la solución única del sistema.

FORMA ESCALONADA. VARIABLES LIBRES

Un sistema de ecuaciones lineales está en *forma escalonada* si ninguna ecuación es degenerada y la primera incógnita de cada ecuación está a la derecha de la primera incógnita de la ecuación anterior. El paradigma es:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= b_1 \\
a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &= b_r
\end{aligned} \qquad [1.5]
$$

donde $1 < j_2 < \cdots < j_r$ y donde $a_{11} \neq 0$, $a_{2j_2} \neq 0, \ldots, a_{rj_r} \neq 0$. Nótese que $r \leqslant n$.

Una incógnita x_k en el sistema escalonado anterior [1.5] se denomina *variable libre* si x_k no es la primera incógnita de ninguna ecuación, esto es, si $x_k \neq x_1$, $x_k \neq x_{j_2}, \ldots, x_k \neq x_{j_r}$.

El siguiente teorema, demostrado en el Problema 1.13, describe el conjunto solución de un sistema escalonado.

Teorema 1.6: Consideremos el sistema de ecuaciones lineales en forma escalonada [1.5]. Existen dos casos:

i) $r = n$. Hay tantas ecuaciones como incógnitas. Entonces el sistema tiene solución única.

ii) $r < n$. Hay menos ecuaciones que incógnitas. Entonces podemos asignar arbitrariamente valores a las $n - r$ variables libres y obtener una solución del sistema.

Supongamos que el sistema escalonado [1.5] contiene más incógnitas que ecuaciones. Entonces el sistema tiene un número infinito de soluciones, porque a cada una de las $n - r$ variables libres se le puede asignar cualquier número real. La solución general del sistema se obtiene como sigue. Se asignan valores arbitrarios, digamos $t_1, t_2, \ldots, t_{n-r}$, llamados *parámetros*, a las variables libres y entonces se emplea la sustitución hacia atrás para obtener los valores de las variables no libres en términos de los parámetros. Alternativamente se puede utilizar la sustitución hacia atrás para despejar directamente las variables no libres $x_1, x_{j_2}, \ldots, x_{j_r}$ en términos de las libres.

EJEMPLO 1.11. Consideremos el sistema

$$
\begin{aligned}
x + 4y - 3z + 2t &= 5 \\
z - 4t &= 2
\end{aligned}
$$

El sistema está en forma escalonada. Las primeras incógnitas son x y z; en consecuencia, las variables libres son las otras incógnitas, y y t.

Para encontrar la solución general del sistema, asignamos valores arbitrarios a las variables libres y y t, digamos $y = a$ y $t = b$, y a continuación utilizamos la sustitución hacia atrás para despejar las variables no libres x y z. Sustituyendo en la última ecuación se obtiene $z - 4b = 2$ o $z = 2 + 4b$. Sustituimos en la primera ecuación para llegar a

$$x + 4a - 3(2 + 4b) + 2b = 5 \quad \text{o} \quad x + 4a - 6 - 12b + 2b = 5 \quad \text{o} \quad x = 11 - 4a + 10b$$

Entonces

$$x = 11 - 4a + 10b,\ y = a,\ z = 2 + 4b,\ t = b \quad \text{o} \quad (11 - 4a + 10b,\ a,\ 2 + 4b,\ b)$$

es la solución general en forma paramétrica. Alternativamente podemos usar la sustitución hacia atrás para despejar las variables no libres x y z, directamente, en términos de las variables libres y y t. La última ecuación da $z = 2 + 4t$. Sustituimos en la primera para obtener

$$x + 4y - 3(2 + 4t) + 2t = 5 \quad \text{o} \quad x + 4y - 6 - 12t + 2t = 5 \quad \text{o} \quad x = 11 - 4y + 10t$$

De acuerdo con esto,

$$\begin{aligned} x &= 11 - 4y + 10t \\ z &= \ \ 2 + 4t \end{aligned}$$

es otra forma para la solución general del sistema.

1.6. ALGORITMO DE REDUCCION

El siguiente algoritmo (a veces llamado reducción por filas) reduce el sistema [1.3] de m ecuaciones lineales con n incógnitas a forma escalonada (posiblemente triangular), o bien determina que el sistema no tiene solución.

ALGORITMO DE REDUCCION

Paso 1. Intercambiar las ecuaciones de forma que x_1 aparezca con un coeficiente no nulo en la primera ecuación; es decir, conseguir que $a_{11} \neq 0$.

Paso 2. Utilizar a_{11} como pivote para eliminar x_1 de todas las ecuaciones excepto de la primera. Esto es, para cada $i > 1$, efectuar la operación (Sección 1.4)

$$[E_3]\text{: } -(a_{i1}/a_{11})L_1 + L_i \to L_i \qquad \text{o} \qquad [E]\text{: } -a_{i1}L_1 + a_{11}L_i \to L_i$$

Paso 3. Examinar cada nueva ecuación L:

a) Si L tiene la forma $0x_1 + 0x_2 + \cdots + 0x_n = 0$ o si es un múltiplo de otra ecuación, suprimirla del sistema.

b) Si L tiene la forma $0x_1 + 0x_2 + \cdots + 0x_n = b$ con $b \neq 0$, abandonar el algoritmo. El sistema no tiene solución.

Paso 4. Repetir los Pasos 1, 2 y 3 con el subsistema formado por todas las ecuaciones, excluyendo la primera.

Paso 5. Continuar el proceso anterior hasta que el sistema esté en forma escalonada, o hasta que se obtenga una ecuación degenerada en el Paso 3 *b*).

La justificación del Paso 3 está en el Teorema 1.5 y en el hecho de que si $L = kL'$ para alguna otra ecuación L' en el sistema, la operación $-kL' + L \to L$ sustituye L por $0x_1 + 0x_2 + \cdots + 0x_n = 0$, que podrá ser de nuevo eliminada según el Teorema 1.5.

EJEMPLO 1.12

a) El sistema

$$\begin{aligned} 2x + y - 2z &= 10 \\ 3x + 2y + 2z &= 1 \\ 5x + 4y + 3z &= 4 \end{aligned}$$

se resuelve reduciéndolo primero a forma escalonada. Para eliminar x de las ecuaciones segunda y tercera efectuamos las operaciones $-3L_1 + 2L_2 \to L_2$ y $-5L_1 + 2L_3 \to L_3$:

$$\begin{array}{rr} -3L_1: & -6x - 3y + 6z = -30 \\ 2L_2: & 6x + 4y + 4z = 2 \\ \hline -3L_1 + 2L_2: & y + 10z = -28 \end{array} \qquad \begin{array}{rr} -5L_1: & -10x - 5y + 10z = -50 \\ 2L_3: & 10x + 8y + 6z = 8 \\ \hline -5L_1 + 2L_3: & 3y + 16z = -42 \end{array}$$

Esto proporciona el siguiente sistema, en el que y se elimina de la tercera ecuación mediante la operación $-3L_2 + L_3 \to L_3$:

$$\left.\begin{aligned} 2x + y - 2z &= 10 \\ y + 10z &= -28 \\ 3y + 16z &= -42 \end{aligned}\right\} \to \left\{\begin{aligned} 2x + y - 2z &= 10 \\ y + 10z &= -28 \\ -14z &= 42 \end{aligned}\right.$$

El sistema está ahora en forma triangular. Por consiguiente, podemos utilizar la sustitución hacia atrás para obtener la solución única $u = (1, 2, -3)$.

b) El sistema

$$\begin{aligned} x + 2y - 3z &= 1 \\ 2x + 5y - 8z &= 4 \\ 3x + 8y - 13z &= 7 \end{aligned}$$

se resuelve reduciéndolo primero a forma escalonada. Para eliminar x de las ecuaciones segunda y tercera efectuamos $-2L_1 + L_2 \to L_2$ y $-3L_1 + L_3 \to L_3$ para obtener

$$\begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \\ 2y - 4z &= 4 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \end{aligned}$$

(La tercera ecuación se suprime puesto que es un múltiplo de la segunda.) El sistema está ahora en forma escalonada con variable libre z.

Para obtener la solución general hacemos $z = a$ y resolvemos por sustitución hacia atrás. Sustituimos $z = a$ en la segunda ecuación para obtener $y = 2 + 2a$. A continuación sustituimos $z = a$ e $y = 2 + 2a$ en la primera ecuación para obtener $x + 2(2 + 2a) - 3a = 1$ o $x = -3 - a$. Entonces la solución general es

$$x = -3 - a,\ y = 2 + 2a,\ z = a \qquad \text{o} \qquad (-3 - a, 2 + 2a, a)$$

donde a es el parámetro.

c) El sistema

$$\begin{aligned} x + 2y - 3z &= -1 \\ 3x - y + 2z &= 7 \\ 5x + 3y - 4z &= 2 \end{aligned}$$

se resuelve reduciéndolo primero a forma escalonada. Para eliminar x de las ecuaciones segunda y tercera efectuamos las operaciones $-3L_1 + L_2 \rightarrow L_2$ y $-5L_1 + L_3 \rightarrow L_3$ para obtener el sistema equivalente

$$\begin{aligned} x + 2y - 3z &= -1 \\ -7y + 11z &= 10 \\ -7y + 11z &= 7 \end{aligned}$$

La operación $-L_2 + L_3 \rightarrow L_3$ conduce a la ecuación degenerada

$$0x + 0y + 0z = -3$$

Entonces el sistema no tiene solución.

El siguiente resultado básico se mencionó previamente.

Teorema 1.7: Cualquier sistema de ecuaciones lineales tiene: i) una única solución, ii) ninguna solución, o iii) un número infinito de soluciones.

Demostración. Aplicando el algoritmo anterior al sistema podemos bien reducirlo a forma escalonada, o bien determinar que no tiene solución. Si la forma escalonada tiene variables libres, entonces el sistema tiene un número infinito de soluciones.

Nota: Se dice que un sistema es *compatible* si tiene una o más soluciones [Casos i) o iii) en el Teorema 1.7], y se dice que es *incompatible* si no tiene solución [Caso ii) en el Teorema 1.7]. La Figura 1-3 ilustra esta situación.

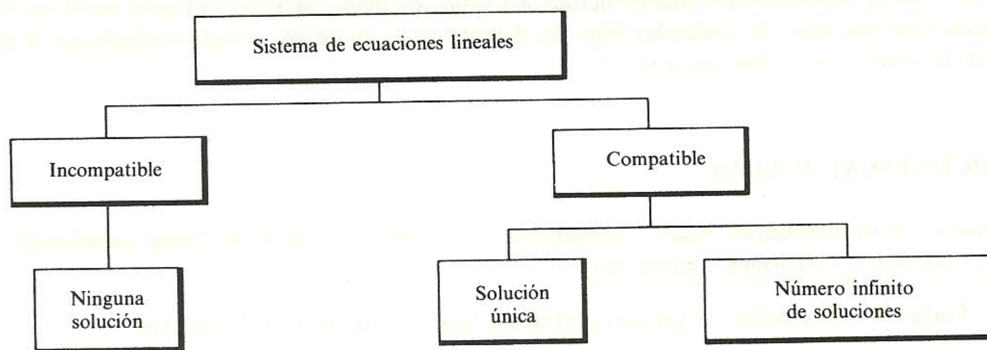


Figura 1-3.

1.7. MATRICES

Sea A una tabla ordenada de números como sigue:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

La tabla A se denomina *matriz*. Tal matriz se denota por $A = (a_{ij})$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, o simplemente $A = (a_{ij})$. Las m n-plas horizontales

$$(a_{11}, a_{12}, \ldots, a_{1n}), (a_{21}, a_{22}, \ldots, a_{2n}), \ldots, (a_{m1}, a_{m2}, \ldots, a_{mn})$$

son las *filas* de la matriz, y las n m-plas verticales

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix}$$

con sus *columnas*. Nótese que el elemento a_{ij}, llamado la *entrada ij* o la *componente ij*, aparece en la fila i-ésima y en la columna j-ésima. Una matriz con m filas y n columnas se llama matriz m por n, o matriz $m \times n$; el par de números (m, n) se llama su *tamaño*.

EJEMPLO 1.13. Sea $A = \begin{pmatrix} 1 & -3 & 4 \\ 0 & 5 & -2 \end{pmatrix}$. Entonces A es una matriz 2×3. Sus filas son $(1, -3, 4)$ y $(0, 5, -2)$; sus columnas son $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -3 \\ 3 \end{pmatrix}$ y $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

La primera entrada no nula en una fila R de una matriz A se llama la entrada *principal* no nula de R. Si R no tiene entrada principal no nula, es decir, si toda entrada en R es 0, R se denomina una *fila nula*. Si todas las filas de A son nulas, es decir, si toda entrada en A es 0, A se llama la *matriz cero*, denotada por 0.

MATRICES ESCALONADAS

Una matriz A se denomina *matriz escalonada*, o se dice que está en *forma escalonada*, si se cumplen las dos condiciones siguientes:

i) Todas las filas nulas, si las hay, están en la parte inferior de la matriz.

ii) Cada entrada principal no nula está a la derecha de la entrada principal no nula de la fila precedente.

Esto es, $A = (a_{ij})$ es una matriz escalonada si existen entradas distintas de cero

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}, \qquad \text{donde} \qquad j_1 < j_2 < \cdots < j_r$$

con la propiedad de que

$$a_{ij} = 0 \qquad \text{para} \qquad i \leqslant r, j < j_i, \qquad \text{y para} \qquad i > r$$

En este caso, $a_{1j_1}, \ldots, a_{rj_r}$ son las entradas principales no nulas de A.

EJEMPLO 1.14. Las siguientes son matrices escalonadas cuyas entradas principales no nulas se han rodeado con un círculo:

$$\begin{pmatrix} \textcircled{2} & 3 & 2 & 0 & 4 & 5 & -6 \\ 0 & 0 & \textcircled{1} & 1 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{6} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \textcircled{1} & 2 & 3 \\ 0 & 0 & \textcircled{1} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \textcircled{1} & 3 & 0 & 0 & 4 \\ 0 & 0 & 0 & \textcircled{1} & 0 & -3 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 2 \end{pmatrix}$$

Se dice que una matriz escalonada A se ha puesto en *forma canónica por filas* si tiene las dos propiedades adicionales siguientes:

iii) Cada entrada principal no nula es 1.

iv) Cada entrada principal no nula es la única entrada distinta de cero en su columna.

La tercera matriz de arriba es un ejemplo de matriz en forma canónica por filas. La segunda no está en dicha forma porque la entrada principal no nula en la segunda fila no es la única entrada distinta de cero en su columna; hay un 3 sobre ella. La primera matriz tampoco está en forma canónica por filas puesto que algunas de las entradas principales no nulas no valen 1.

La matriz cero, 0, con cualquier número de filas y de columnas, es otro ejemplo de matriz en forma canónica por filas.

1.8. EQUIVALENCIA POR FILAS. OPERACIONES ELEMENTALES ENTRE FILAS

Se dice que una matriz A es *equivalente por filas* a otra B, escrito $A \sim B$, si B puede obtenerse a partir de A mediante una sucesión finita de las siguientes operaciones, llamadas *operaciones elementales entre filas*:

[E_1] Intercambiar las filas i-ésima y j-ésima: $R_i \leftrightarrow R_j$.

[E_2] Multiplicar la fila i-ésima por un escalar no nulo k: $kR_i \to R_i$, $k \neq 0$.

[E_3] Sustituir la fila i-ésima por ella misma más k veces la j-ésima: $kR_j + R_i \to R_i$.

En la práctica efectuamos [E_2] y [E_3] en un solo paso, es decir, la operación

[E] Sustituir la fila i-ésima por k (no nulo) veces ella misma más k' veces la j-ésima:

$$k'R_j + kR_i \to R_i, \ k \neq 0$$

Sin duda, el lector reconocerá la similitud entre las operaciones anteriores y las utilizadas para resolver sistemas de ecuaciones lineales.

El siguiente algoritmo reduce por filas una matriz A a forma escalonada. (La expresión «reducir por filas» o, simplemente, «reducir» significará transformar una matriz mediante operaciones elementales entre filas.)

Algoritmo 1.8A

Aquí $A = (a_{ij})$ es una matriz arbitraria.

Paso 1. Encontrar la primera columna con una entrada no nula. Supongamos que es la columna j_1.

Paso 2. Intercambiar las filas de forma que aparezca una entrada no nula en la primera fila de la columna j_1, esto es, conseguir que $a_{1j_1} \neq 0$.

Paso 3. Utilizar a_{1j_1} como pivote para obtener cero bajo él; esto es, para cada $i > 1$ efectuar la operación entre filas $-a_{ij_1}R_1 + a_{1j_1}R_i \to R_i$ o $(-a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$.

Paso 4. Repetir los Pasos 1, 2 y 3 con la submatriz formada por todas las filas, excluyendo la primera.

Paso 5. Continuar el proceso anterior hasta que la matriz quede en forma escalonada.

EJEMPLO 1.15. La matriz $A = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{pmatrix}$ se reduce a forma escalonada mediante el Algoritmo 1.8A como sigue:

Utilizamos $a_{11} = 1$ como pivote para obtener ceros bajo a_{11}, esto es, efectuamos las operaciones entre filas $-2R_1 + R_2 \to R_2$ y $-3R_1 + R_3 \to R_3$ para obtener la matriz

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 3 \end{pmatrix}$$

Ahora utilizamos $a_{23} = 4$ como pivote para obtener un cero bajo a_{23}, esto es, efectuamos la operación entre filas $-5R_2 + 4R_3 \to R_3$, obteniendo la matriz

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

La matriz está ahora en forma escalonada.

El siguiente algoritmo reduce por filas una matriz escalonada a su forma canónica por filas.

Algoritmo 1.8B

Aquí $A = (a_{ij})$ está en forma escalonada, digamos con entradas principales no nulas

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$$

Paso 1. Multiplicar la última fila no nula R_r por $1/a_{rj_r}$ de forma que la entrada principal no nula sea 1.

Paso 2. Utilizar $a_{rj_r} = 1$ como pivote para obtener ceros sobre él; esto es, para $i = r - 1$, $r - 2, \ldots, 1$, efectuar la operación

$$-a_{ir_i}R_r + R_i \to R_i$$

Paso 3. Repetir los Pasos 1 y 2 para las filas $R_{r-1}, R_{r-2}, \ldots, R_2$.

Paso 4. Multiplicar R_1 por $1/a_{1j_1}$.

EJEMPLO 1.16. Utilizando el Algoritmo 1.8B, la matriz escalonada

$$A = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 3 & 2 & 5 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

se reduce a forma canónica por filas como sigue:

Multiplicamos R_3 por $\frac{1}{4}$ de forma que su entrada principal no nula sea 1; entonces utilizamos $a_{35} = 1$ como pivote para obtener ceros sobre él efectuando las operaciones $-5R_3 + R_2 \rightarrow R_2$ y $-6R_3 + R_1 \rightarrow R_1$:

$$A \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 3 & 2 & 5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Multiplicamos R_2 por $\frac{1}{3}$ de modo que su entrada principal no nula sea 1; entonces utilizamos $a_{23} = 1$ como pivote para conseguir un 0 encima, con la operación $-4R_2 + R_1 \rightarrow R_1$:

$$A \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 0 \\ 0 & 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 3 & 0 & \frac{7}{3} & 0 \\ 0 & 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Finalmente, multiplicamos R_1 por $\frac{1}{2}$ obteniendo

$$\begin{pmatrix} 1 & \frac{3}{2} & 0 & \frac{7}{6} & 0 \\ 0 & 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Esta matriz es la forma canónica por filas de A.

Los Algoritmos 1.8A y 1.8B muestran que cualquier matriz es equivalente por filas a al menos una matriz en forma canónica por filas. En el Capítulo 5 demostraremos que tal matriz es única, esto es:

Teorema 1.8: Cualquier matriz A es equivalente por filas a una única matriz en forma canónica por filas (llamada la *forma canónica por filas* de A).

Nota: Si una matriz A está en forma escalonada, sus entradas principales no nulas se denominarán *entradas pivote*. El término proviene del algoritmo anterior que reduce por filas una matriz a forma escalonada.

1.9. SISTEMAS DE ECUACIONES LINEALES Y MATRICES

La matriz ampliada M del sistema [1.3] de m ecuaciones lineales con n incógnitas es la siguiente:

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Obsérvese que cada fila de M corresponde a una ecuación del sistema y cada columna a los coeficientes de una incógnita, excepto la última, que corresponde a las constantes del sistema.

La matriz de coeficientes A del sistema [1.3] es

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Nótese que la matriz de coeficientes A puede obtenerse de la ampliada M omitiendo su última columna.

Un sistema de ecuaciones lineales puede resolverse trabajando con su matriz ampliada. Específicamente, reduciéndola a forma escalonada (lo que nos dice si el sistema es compatible) y luego a su forma canónica por filas. La justificación de este proceso proviene de los siguientes hechos:

1. Cualquier operación elemental entre filas en la matriz ampliada M del sistema es equivalente a efectuar la operación correspondiente en el sistema mismo.
2. El sistema tiene solución si y sólo si la forma escalonada de la matriz ampliada no tiene una fila de la forma $(0, 0, \dots, 0, b)$ con $b \neq 0$.
3. En la forma canónica por filas de la matriz ampliada M (excluyendo filas nulas) el coeficiente de cada variable no libre es una entrada principal no nula igual a uno y es la única entrada distinta de cero en su columna; de aquí la solución en forma de variables libres se obtiene simplemente transfiriendo los términos de variable no libre al otro miembro.

Este proceso se ilustra en el siguiente ejemplo.

EJEMPLO 1.17

a) El sistema

$$\begin{aligned} x + y - 2z + 4t &= 5 \\ 2x + 2y - 3z + t &= 3 \\ 3x + 3y - 4z - 2t &= 1 \end{aligned}$$

se resuelve reduciendo su matriz ampliada M a forma escalonada y después a forma canónica por filas como sigue:

$$M = \begin{pmatrix} 1 & 1 & -2 & 4 & 5 \\ 2 & 2 & -3 & 1 & 3 \\ 3 & 3 & -4 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \end{pmatrix}$$

[La tercera fila (en la segunda matriz) se suprime, puesto que es un múltiplo de la segunda y resultará una fila nula.] Así la solución general en forma de variables libres del sistema es como se indica a continuación:

$$\begin{aligned} x + y \qquad - 10t &= -9 \\ z - 7t &= -7 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x &= -9 - y + 10t \\ z &= -7 + 7t \end{aligned}$$

Aquí las variables libres son y y t, y las no libres x y z.

b) El sistema

$$\begin{aligned} x_1 + x_2 - 2x_3 + 3x_4 &= 4 \\ 2x_1 + 3x_2 + 3x_3 - x_4 &= 3 \\ 5x_1 + 7x_2 + 4x_3 + x_4 &= 5 \end{aligned}$$

se resuelve como sigue. Para empezar, reducimos su matriz ampliada a forma escalonada:

$$M = \begin{pmatrix} 1 & 1 & -2 & 3 & 4 \\ 2 & 3 & 3 & -1 & 3 \\ 5 & 7 & 4 & 1 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 2 & 14 & -14 & 15 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix}$$

No hay necesidad de continuar para hallar la forma canónica por filas de la matriz, puesto que la matriz escalonada ya nos indica que el sistema no tiene solución. Específicamente, la tercera fila en la matriz escalonada corresponde a la ecuación degenerada

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -5$$

que no tiene solución.

c) El sistema

$$\begin{aligned} x + 2y + z &= 3 \\ 2x + 5y - z &= -4 \\ 3x - 2y - z &= 5 \end{aligned}$$

se resuelve reduciendo su matriz ampliada a forma escalonada y luego a forma canónica por filas como sigue:

$$M = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & -1 & -4 \\ 3 & -2 & -1 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & -8 & -4 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & -28 & -84 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

De este modo, el sistema tiene la solución única $x = 2$, $y = -1$, $z = 3$ o $u = (2, -1, 3)$. (Nótese que la forma escalonada de M ya indicaba que la solución era única, puesto que correspondía a un sistema triangular.)

1.10. SISTEMAS DE ECUACIONES LINEALES HOMOGENEOS

Se dice que el sistema de ecuaciones lineales [1.3] es *homogéneo* si todas las constantes son iguales a cero, esto es, si tiene la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \qquad [1.6]$$

De hecho, [1.6] se denomina el sistema homogéneo asociado con el sistema [1.3].

El sistema homogéneo [1.6] siempre tiene una solución, a saber, la n-pla nula $0 = (0, 0, \ldots, 0)$ llamada la solución nula o trivial. (Cualquier otra solución, si existe, se denomina solución no nula o no trivial.) Siendo así, el sistema siempre puede reducirse a uno homogéneo equivalente en forma escalonada:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= 0 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &= 0 \end{aligned} \qquad [1.7]$$

Existen dos posibilidades:

i) $r = n$. En tal caso, el sistema tiene sólo la solución nula.

ii) $r < n$. Entonces el sistema tiene una solución no nula.

De acuerdo con esto, si empezamos con menos ecuaciones que incógnitas, tendremos, en forma escalonada, $r < n$ y por consiguiente el sistema tendrá una solución no nula. Esto prueba el importante teorema siguiente.

Teorema 1.9: Un sistema de ecuaciones lineales homogéneo con más incógnitas que ecuaciones tiene solución no nula.

EJEMPLO 1.18

a) El sistema homogéneo

$$\begin{aligned} x + 2y - 3z + w &= 0 \\ x - 3y + z - 2w &= 0 \\ 2x + y - 3z + 5w &= 0 \end{aligned}$$

tiene una solución no nula porque hay cuatro incógnitas pero sólo tres ecuaciones.

b) Reducimos el siguiente sistema a forma escalonada:

$$\begin{aligned} x + y - z &= 0 \\ 2x - 3y + z &= 0 \\ x - 4y + 2z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ -5y + 3z &= 0 \\ -5y + 3z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ -5y + 3z &= 0 \end{aligned}$$

El sistema tiene una solución no nula, puesto que hemos obtenido sólo dos ecuaciones para las tres incógnitas en forma escalonada. Por ejemplo, sea $z = 5$; entonces $y = 3$ y $x = 2$. Dicho de otro modo, la terna (2, 3, 5) es una solución particular no nula.

c) Reducimos el siguiente sistema a forma escalonada:

$$\begin{aligned} x + y - z &= 0 \\ 2x + 4y - z &= 0 \\ 3x + 2y + 2z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ 2y + z &= 0 \\ -y + 5z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ 2y + z &= 0 \\ 11z &= 0 \end{aligned}$$

Como en forma escalonada hay tres ecuaciones y tres incógnitas, el sistema dado tiene sólo la solución nula (0, 0, 0).

BASE PARA LA SOLUCION GENERAL DE UN SISTEMA HOMOGENEO

Denotemos por W la solución general de un sistema homogéneo. Se dice que los vectores solución no nulos $u_1, u_2, \ldots, u_s$ forman una *base* de W si todo vector solución w en W puede expresarse de forma única como combinación lineal de $u_1, u_2, \ldots, u_s$. El número s de tales vectores base se denomina la *dimensión* de W, escrito dim $W = s$. (Si $W = \{0\}$, dim $W = 0$, por definición.)

El siguiente teorema, probado en el Capítulo 5, nos dice cómo hallar tal base.

Teorema 1.10: Sea W la solución general de un sistema homogéneo y supongamos que la forma escalonada del sistema tiene s variables libres. Sean $u_1, u_2, \ldots, u_s$ las soluciones obtenidas haciendo una de las variables libres igual a uno (o a cualquier constante distinta de cero) y las restantes variables libres iguales a cero. Entonces dim $W = s$ y $u_1, u_2, \ldots, u_s$ forman una base de W.

Nota: La expresión *combinación lineal* utilizada anteriormente se refiere a suma de productos de un vector por un escalar, donde tales operaciones se definen según

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$
$$k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$$

Estas operaciones se estudiarán en detalle en el Capítulo 2.

EJEMPLO 1.19. Supongamos que queremos encontrar la dimensión y una base para la solución general W del sistema homogéneo

$$\begin{aligned} x + 2y - 3z + 2s - 4t &= 0 \\ 2x + 4y - 5z + s - 6t &= 0 \\ 5x + 10y - 13z + 4s - 16t &= 0 \end{aligned}$$

Para empezar reducimos el sistema a forma escalonada. Efectuando las operaciones $-2L_1 + L_2 \to L_2$ y $-5L_2 + L_3 \to L_3$ y después $-2L_2 + L_3 \to L_3$ se obtiene:

$$\begin{aligned} x + 2y - 3z + 2s - 4t &= 0 \\ z - 3s + 2t &= 0 \\ 2z - 6s + 4t &= 0 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x + 2y - 3z + 2s - 4t &= 0 \\ z - 3s + 2t &= 0 \end{aligned}$$

En forma escalonada, el sistema tiene tres variables libres, y, s y t; de aquí dim $W = 3$. Se obtienen tres vectores solución que forman base de W como sigue:

1. Hacemos $y = 1$, $s = 0$, $t = 0$. La sustitución hacia atrás proporciona la solución $u_1 = (-2, 1, 0, 0, 0)$.
2. Hacemos $y = 0$, $s = 1$, $t = 0$. La sustitución hacia atrás proporciona la solución $u_2 = (7, 0, 3, 1, 0)$.
3. Hacemos $y = 0$, $s = 0$, $t = 1$. La sustitución hacia atrás proporciona la solución $u_3 = (-2, 0, -2, 0, 1)$.

El conjunto $\{u_1, u_2, u_3\}$ es una base de W.

Ahora bien, cualquier solución del sistema puede escribirse de la forma

$$au_1 + bu_2 + cu_3 = a(-2, 1, 0, 0, 0) + b(7, 0, 3, 1, 0) + c(-2, 0, -2, 0, 1) = (-2a + 7b - 2c, a, 3b - 2c, b, c)$$

donde a, b y c son constantes arbitrarias. Obsérvese que esto no es más que la forma paramétrica de la solución general con la elección de parámetros $y = a$, $s = b$ y $t = c$.

SISTEMAS INHOMOGENEOS Y SUS SISTEMAS HOMOGENEOS ASOCIADOS

La relación entre el sistema inhomogéneo [1.3] y su sistema homogéneo asociado [1.6] está contenida en el siguiente teorema, cuya demostración se pospone hasta el Capítulo 3 (Teorema 3.5).

Teorema 1.11: Sean v una solución particular y U la solución general de un sistema inhomogéneo de ecuaciones lineales. En tal caso,

$$U = v_0 + W = \{v_0 + w : w \in W\}$$

donde W es la solución general del sistema homogéneo asociado.

Esto es, $U = v_0 + W$ puede obtenerse sumando v_0 a cada elemento de W.

El teorema anterior tiene una interpretación geométrica en el espacio $\mathbf{R}^3$. De forma específica, si W es una recta que pasa por el origen, como se muestra en la Figura 1-4, $U = v_0 + W$ es la recta paralela a W que puede obtenerse sumando v_0 a cada elemento de W. Análogamente, siempre que W sea un plano por el origen, entonces $U = v_0 + W$ es un plano paralelo a W.

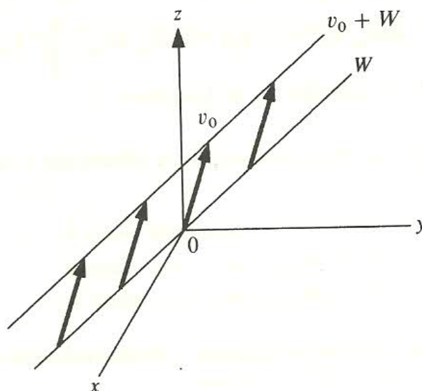


Figura 1-4.

PROBLEMAS RESUELTOS

ECUACIONES LINEALES. SOLUCIONES

1.1. Determinar si cada una de las siguientes ecuaciones es lineal:

a) $5x + 7y - 8yz = 16$ *b*) $x + \pi y + ez = \log 5$ *c*) $3x + ky - 8z = 16$

a) No, porque el producto yz de dos incógnitas es de segundo grado.

b) Sí, puesto que π, e y $\log 5$ son constantes.

c) Tal y como aparece, hay cuatro incógnitas: x, y, z, k. Debido al término ky, no es una ecuación lineal. No obstante, supuesto que k sea una constante, la ecuación es lineal en las incógnitas x, y, z.

1.2. Considérese la ecuación lineal $x + 2y - 3z = 4$. Determinar si $u = (8, 1, 2)$ es solución.

Como el orden de las incógnitas es x, y, z, $u = (8, 1, 2)$ es una abreviatura de $x = 8$, $y = 1$, $z = 2$. Sustituimos en la ecuación para obtener

$$8 + 2(1) - 3(2) = 4 \quad \text{o} \quad 8 + 2 - 6 = 4 \quad \text{o} \quad 4 = 4$$

Sí, es una solución.

1.3. Determinar si *a*) $u = (3, 2, 1, 0)$ y *b*) $v = (1, 2, 4, 5)$ son soluciones de la ecuación $x_1 + 2x_2 - 4x_3 + x_4 = 3$.

a) Sustituimos para obtener $3 + 2(2) - 4(1) + 0 = 3$, o $3 = 3$; sí, es una solución.

b) Sustituimos para obtener $1 + 2(2) - 4(4) + 5 = 3$, o $-6 = 3$; no es solución.

1.4. ¿Es $u = (6, 4, -2)$ una solución de la ecuación $3x_2 + x_3 - x_1 = 4$?

Por convenio, las componentes de u están ordenadas de acuerdo con los subíndices de las incógnitas. Esto es, $u = (6, 4, -2)$ es una abreviatura de $x_1 = 6$, $x_2 = 4$, $x_3 = -2$. Sustituyendo en la ecuación obtenemos $3(4) - 2 - 6 = 4$, o $4 = 4$. Sí, es una solución.

1.5. Probar el Teorema 1.1.

Supongamos $a \neq 0$. Entonces existe el escalar b/a. Sustituyendo b/a en $ax = b$ se obtiene $a(b/a) = b$, o $b = b$; por consiguiente, b/a es una solución. Por otra parte, supongamos que x_0 es solución de $ax = b$, de forma que $ax_0 = b$. Multiplicando ambos miembros por $1/a$ se obtiene $x_0 = b/a$. De aquí b/a es la única solución de $ax = b$. Por tanto, i) queda demostrado.

Sea ahora, por el contrario, $a = 0$. Entonces, para todo escalar k, tenemos $ak = 0k = 0$. Si $b \neq 0$, $ax \neq b$. De acuerdo con esto, k no es solución de $ax = b$ y queda así demostrado ii). Si $b = 0$, $ak = b$. Esto es, cualquier escalar k es una solución de $ax = b$ y queda demostrado iii).

1.6. Resolver cada una de las siguientes ecuaciones:

a) $ex = \log 5$ *c*) $3x - 4 - x = 2x + 3$

b) $cx = 0$ *d*) $7 + 2x - 4 = 3x + 3 - x$

a) Como $e \neq 0$, multiplicamos por $1/e$ obteniendo $x = (\log 5)/e$.

b) Si $c \neq 0$, $0/c = 0$ es la única solución. Si $c = 0$, todo escalar k es solución [Teorema 1.1 iii)].

c) La reescribimos en forma convencional, $2x - 4 = 2x + 3$ ó $0x = 7$. La ecuación no tiene solución [Teorema 1.1 ii)].

d) La reescribimos en forma convencional, $3 + 2x = 2x + 3$ ó $0x = 0$. Todo escalar k es una solución [Teorema 1.1 iii)].

1.7. Describir las soluciones de la ecuación $2x + y + x - 5 = 2y + 3x - y + 4$.

La reescribimos en forma convencional, agrupando términos y trasponiendo:

$$3x + y - 5 = y + 3x + 4 \quad \text{o} \quad 0x + 0y = 9$$

La ecuación es degenerada con una constante no nula. Siendo así, la ecuación no tiene solución.

1.8. Describir las soluciones de la ecuación $2y + 3x - y + 4 = x + 3 + y + 1 + 2x$.

La reescribimos en forma convencional, agrupando términos y trasponiendo:

$$y + 3x + 4 = 3x + 4 + y \qquad \text{o} \qquad 0x + 0y = 0$$

La ecuación es degenerada con constante nula; en consecuencia, todo vector $u = (a, b)$ en $\mathbf{R}^2$ es una solución.

1.9. Probar el Teorema 1.3.

Probemos primero i). Tomamos $x_j = k_j$ para $j \neq p$. Como $a_j = 0$ para $j < p$, la sustitución en la ecuación nos conduce a

$$a_p x_p + a_{p+1}k_{p+1} + \cdots + a_n k_n = b \qquad \text{o} \qquad a_p x_p = b - a_{p+1}k_{p+1} - \cdots - a_n k_n$$

con $a_p \neq 0$. Por el Teorema 1.1 i), x_p está unívocamente determinado como:

$$x_p = \frac{1}{a_p}(b - a_{p+1}k_{p+1} - \cdots - a_n k_n)$$

Por tanto, queda probado i).

Probemos ahora ii). Supongamos que $u = (k_1, k_2, \ldots, k_n)$ es una solución. Entonces

$$a_p k_p + a_{p+1}k_{p+1} + \cdots + a_n k_n = b \qquad \text{o} \qquad k_p = \frac{1}{a_p}(b - a_{p+1}k_{p+1} - \cdots - a_n k_n)$$

De cualquier modo, ésta es precisamente la solución

$$u = \left(k_1, \ldots, k_{p-1}, \frac{b - a_{p+1}k_{p+1} - \cdots - a_n k_n}{a_p}, k_{p+1}, \ldots, k_n\right)$$

obtenida en i). Queda así probado ii).

1.10. Considérese la ecuación lineal $x - 2y + 3z = 4$. Hallar: *a*) tres soluciones particulares, *b*) la solución general.

a) Aquí x es la primera incógnita. De acuerdo con ello, asignamos valores cualesquiera a las variables libres y y z y luego despejamos x para obtener una solución. Por ejemplo:

1. Tomamos $y = 1$ y $z = 1$. La sustitución en la ecuación nos conduce a

$$x - 2(1) + 3(1) = 4 \qquad \text{o} \qquad x - 2 + 3 = 4 \qquad \text{o} \qquad x = 3$$

Por consiguiente, $u_1 = (3, 1, 1)$ es una solución.

2. Tomamos $y = 1$, $z = 0$. La sustitución proporciona $x = 6$; por tanto, $u_2 = (6, 1, 0)$ es una solución.

3. Tomamos $y = 0$, $z = 1$. La sustitución proporciona $x = 1$; por tanto, $u_3 = (1, 0, 1)$ es una solución.

b) Para hallar la solución general asignamos valores arbitrarios a las variables libres, digamos $y = a$ y $z = b$. (Llamamos a a y b parámetros de la solución.) Después sustituimos en la ecuación para obtener

$$x - 2a + 3b = 4 \qquad \text{o} \qquad x = 4 + 2a - 3b$$

Así $u = (4 + 2a - 3b, a, b)$ es la solución general.

SISTEMAS EN FORMA TRIANGULAR Y ESCALONADA

1.11. Resolver el sistema

$$\begin{aligned} 2x - 3y + 5z - 2t &= 9 \\ 5y - z + 3t &= 1 \\ 7z - t &= 3 \\ 2t &= 8 \end{aligned}$$

El sistema está en forma triangular; por tanto, resolvemos por sustitución hacia atrás.

i) La última ecuación proporciona $t = 4$.

ii) Sustituyendo en la tercera se obtiene $7z - 4 = 3$, $7z = 7$ o $z = 1$.

iii) Sustituyendo $z = 1$ y $t = 4$ en la segunda ecuación obtenemos

$$5y - 1 + 3(4) = 1 \quad \text{o} \quad 5y - 1 + 12 = 1 \quad \text{o} \quad 5y = -10 \quad \text{o} \quad y = -2$$

iv) Sustituyendo $y = -2$, $z = 1$, $t = 4$ en la primera ecuación tendremos

$$2x - 3(-2) + 5(1) - 2(4) = 9 \quad \text{o} \quad 2x + 6 + 5 - 8 = 9 \quad \text{o} \quad 2x = 6 \quad \text{o} \quad x = 3$$

Así $x = 3$, $y = -2$, $z = 1$, $t = 4$ es la solución única del sistema.

1.12. Determinar las variables libres en cada uno de los sistemas:

$$\begin{aligned} 3x + 2y - 5z - 6s + 2t &= 4 \\ z + 8s - 3t &= 6 \\ s - 5t &= 5 \end{aligned} \qquad \begin{aligned} 5x - 3y + 7z &= 1 \\ 4y + 5z &= 6 \\ 4z &= 9 \end{aligned} \qquad \begin{aligned} x + 2y - 3z &= 2 \\ 2x - 3y + z &= 1 \\ 5x - 4y - z &= 4 \end{aligned}$$

a) *b*) *c*)

a) En forma escalonada, cualquier incógnita que no sea primera es variable libre. Aquí y y t son las variables libres.

b) Las primeras incógnitas son x, y, z. Por consiguiente, no hay variables libres (como en cualquier sistema triangular).

c) La noción de variable libre se aplica sólo a sistemas en forma escalonada.

1.13. Probar el Teorema 1.6.

Existen dos casos:

i) $r = n$. Esto es, hay tantas ecuaciones como incógnitas. En tal caso, el sistema tiene solución única.

ii) $r < n$. Esto es, hay menos ecuaciones que incógnitas. Entonces podemos asignar valores a las $n - r$ variables libres arbitrariamente y obtener una solución del sistema.

Se demuestra por inducción en el número de ecuaciones del sistema, r. Si $r = 1$, tenemos una sola ecuación lineal, no degenerada, a la que se aplica el Teorema 1.3 cuando $n > r = 1$, y el Teorema 1.1 cuando $n = r = 1$. De este modo, el teorema es válido para $r = 1$.

Ahora supongamos que $r > 1$ y que el teorema es válido para un sistema con $r - 1$ ecuaciones. Podemos ver las $r - 1$ ecuaciones

$$\begin{gathered} a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n = b_r \end{gathered}$$

como un sistema con incógnitas $x_{j_2}, \ldots, x_n$. Nótese que el sistema está en forma escalonada. Por la hipótesis de inducción, podemos asignar arbitrariamente valores a las $(n - j_2 + 1) - (r - 1)$ variables libres en el sistema reducido para obtener una solución (digamos $x_{j_2} = k_{j_2}, \ldots, x_n = k_n$). Como en el caso $r = 1$, estos valores, junto con valores arbitrarios de las $j_2 - 2$ variables libres adicionales (digamos $x_2 = k_2, \ldots, x_{j_2-1} = k_{j_2-1}$), conducen a una solución de la primera ecuación con

$$x_1 = \frac{1}{a_{11}} (b_1 - a_{12}k_2 - \cdots - a_{1n}k_n)$$

[Obsérvese que hay $(n - j_2 + 1) - (r - 1) + (j_2 - 2) = n - r$ variables libres.] Además, estos valores para $x_1, \ldots, x_n$ también satisfacen las otras ecuaciones ya que, en éstas, los coeficientes de $x_1, \ldots, x_{j_2-1}$ son nulos.

Ahora, si $r = n$, entonces $j_2 = 2$. Así obtenemos, por inducción, una solución única del subsistema y por ende una solución única del sistema completo. De acuerdo con esto, el teorema queda demostrado.

1.14. Hallar la solución general del sistema escalonado

$$\begin{aligned} x - 2y - 3z + 5s - 2t &= 4 \\ 2z - 6s + 3t &= 2 \\ 5t &= 10 \end{aligned}$$

Como las ecuaciones empiezan con las incógnitas x, z y t, respectivamente, las restantes, y y s, son las variables libres. Para encontrar la solución general asignamos parámetros a las variables libres, digamos $y = a$ y $s = b$, y utilizar sustitución hacia atrás para despejar las variables no libres x, z y t.

i) La última ecuación conduce a $t = 2$.

ii) Sustituimos $t = 2$, $s = b$ en la segunda ecuación para obtener

$$2z - 6b + 3(2) = 2 \quad \text{o} \quad 2z - 6b + 6 = 2 \quad \text{o} \quad 2z = 6b - 4 \quad \text{o} \quad z = 3b - 2$$

iii) Sustituimos $t = 2$, $s = b$, $z = 3b - 2$, $y = a$ en la primera ecuación para obtener

$$x - 2a - 3(3b - 2) + 5b - 2(2) = 4 \quad \text{o} \quad x - 2a - 9b + 6 + 5b - 4 = 4 \quad \text{o} \quad x = 2a + 4b + 2$$

Así

$$x = 2a + 4b + 2 \qquad y = a \qquad z = 3b - 2 \qquad s = b \qquad t = 2$$

o, equivalentemente,

$$u = (2a + 4b + 2,\ a,\ 3b - 2,\ b,\ 2)$$

es la forma paramétrica de la solución general.

Alternativamente, despejar x, z y t en función de las variables libres y y s proporciona la solución general en forma de variables libres:

$$x = 2y + 4s + 2 \qquad z = 3s - 2 \qquad t = 2$$

SISTEMAS DE ECUACIONES LINEALES. ELIMINACION GAUSSIANA

1.15. Resolver el sistema

$$\begin{aligned} x - 2y + z &= 7 \\ 2x - y + 4z &= 17 \\ 3x - 2y + 2z &= 14 \end{aligned}$$

Lo reducimos a forma escalonada. Efectuamos $-2L_1 + L_2 \to L_2$ y $-3L_1 + L_3 \to L_3$ para eliminar x de las ecuaciones segunda y tercera y a continuación efectuamos $-4L_2 + 3L_3 \to L_3$ para eliminar y de la tercera ecuación. Estas operaciones conducen a

$$\begin{aligned} x - 2y + z &= 7 \\ 3y + 2z &= 3 \\ 4y - z &= -7 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x - 2y + z &= 7 \\ 3y + 2z &= 3 \\ -11z &= -33 \end{aligned}$$

El sistema está en forma triangular y por tanto tras la sustitución hacia atrás, tiene la solución única $u = (2, -1, 3)$.

1.16 Resolver el sistema

$$\begin{aligned} 2x - 5y + 3z - 4s + 2t &= 4 \\ 3x - 7y + 2z - 5s + 4t &= 9 \\ 5x - 10y - 5z - 4s + 7t &= 22 \end{aligned}$$

Lo reducimos a forma escalonada. Efectuamos las operaciones $-3L_1 + 2L_2 \to L_2$ y $-5L_1 + 2L_3 \to L_3$ y luego $-5L_2 + L_3 \to L_3$ para obtener

$$\begin{aligned} 2x - 5y + 3z - 4s + 2t &= 4 \\ y - 5z + 2s + 2t &= 6 \\ 5y - 25z + 12s + 4t &= 24 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 2x - 5y + 3z - 4s + 2t &= 4 \\ y - 5z + 2s + 2t &= 6 \\ 2s - 6t &= -6 \end{aligned}$$

El sistema está ahora en forma escalonada. Despejando las primeras incógnitas, x, y y s, en términos de las variables libres, z y t, obtenemos la solución general en forma de variables libres:

$$x = 26 + 11z - 15t \qquad y = 12 + 5z - 8t \qquad s = -3 + 3t$$

De ésta se obtiene de una vez la forma paramétrica de la solución general (con $z = a$, $t = b$):

$$x = 26 + 11a - 15b \qquad y = 12 + 5a - 8b \qquad z = a \qquad s = -3 + 3b \qquad t = b$$

1.17. Resolver el sistema

$$\begin{aligned} x + 2y - 3z + 4t &= 2 \\ 2x + 5y - 2z + t &= 1 \\ 5x + 12y - 7z + 6t &= 7 \end{aligned}$$

Reducimos el sistema a forma escalonada. Eliminamos x de las ecuaciones segunda y tercera mediante las operaciones $-2L_1 + L_2 \to L_2$ y $-5L_1 + L_3 \to L_3$; esto proporciona el sistema

$$\begin{aligned} x + 2y - 3z + 4t &= 2 \\ y + 4z - 7t &= -3 \\ 2y + 8z - 14t &= -3 \end{aligned}$$

La operación $-2L_2 + L_3 \to L_3$ proporciona la ecuación degenerada $0 = 3$. Siendo así, el sistema no tiene solución (incluso a pesar de tener más incógnitas que ecuaciones).

1.18. Determinar los valores de k para que el siguiente sistema, con incógnitas x, y, z, tenga: i) una solución única, ii) ninguna solución, iii) infinitas soluciones.

$$\begin{aligned} x + y - z &= 1 \\ 2x + 3y + kz &= 3 \\ x + ky + 3y &= 2 \end{aligned}$$

Reducimos el sistema a forma escalonada. Eliminamos x de las ecuaciones segunda y tercera mediante las operaciones $-2L_1 + L_2 \to L_2$ y $-L_1 + L_3 \to L_3$ para obtener

$$\begin{aligned} x + \quad y - \quad z &= 1 \\ y + (k+2)z &= 1 \\ (k-1)y + \quad 4z &= 1 \end{aligned}$$

Para eliminar y de la tercera ecuación efectuamos la operación $-(k-1)L_2 + L_3 \to L_3$ para obtener

$$\begin{aligned} x + y - \quad z &= 1 \\ y + \quad (k+2)z &= 1 \\ (3+k)(2-k)z &= 2-k \end{aligned}$$

El sistema tiene solución única si el coeficiente de z en la tercera ecuación es distinto de cero; esto es, si $k \neq 2$ y $k \neq -3$. En el caso $k = 2$, la tercera ecuación se reduce a $0 = 0$ y el sistema tiene un número infinito de soluciones (una para cada valor de z). En el caso $k = -3$, la tercera ecuación se reduce a $0 = 5$ y el sistema no tiene solución. Resumiendo: i) $k \neq 2$ y $k \neq 3$, ii) $k = -3$, iii) $k = 2$.

1.19. ¿Qué condición debe imponerse a a, b y c para que el siguiente sistema, con incógnitas x, y y z, tenga solución?

$$\begin{aligned} x + 2y - \quad 3z &= a \\ 2x + 6y - 11z &= b \\ x - 2y + \quad 7z &= c \end{aligned}$$

Reducimos el sistema a forma escalonada. Eliminando x de las ecuaciones segunda y tercera mediante las operaciones $-2L_1 + L_2 \to L_2$ y $-L_1 + L_3 \to L_3$ obtenemos el sistema equivalente:

$$\begin{aligned} x + 2y - \quad 3z &= a \\ 2y - \quad 5z &= b - 2a \\ -4y + 10z &= c - a \end{aligned}$$

Eliminando y de la tercera ecuación mediante la operación $2L_2 + L_3 \to L_3$ obtenemos finalmente el sistema equivalente

$$\begin{aligned} x + 2y - 3z &= a \\ 2y - 5z &= b - 2a \\ 0 &= c + 2b - 5a \end{aligned}$$

El sistema no tendrá solución si $c + 2b - 5a \neq 0$. Así el sistema tendrá al menos una solución si $c + 2b - 5a = 0$, o $5a = 2b + c$. Obsérvese que en este caso el sistema tendrá infinitas soluciones. En otras palabras, el sistema no puede tener solución única.

MATRICES. MATRICES ESCALONADAS. REDUCCION POR FILAS

1.20. Intercambiar las filas en cada una de las matrices siguientes para obtener una matriz escalonada:

$$\begin{pmatrix} 0 & 1 & -3 & 4 & 6 \\ 4 & 0 & 2 & 5 & -3 \\ 0 & 0 & 7 & -2 & 8 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 5 & -4 & 7 \end{pmatrix} \qquad \begin{pmatrix} 0 & 2 & 2 & 2 & 2 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

a) *b)* *c)*

a) Intercambiamos las filas primera y segunda, es decir, efectuamos la operación elemental entre filas $R_1 \leftrightarrow R_2$.

b) Llevamos la fila nula a la parte inferior de la matriz, es decir, efectuamos la operación $R_1 \leftrightarrow R_2$ y luego $R_2 \leftrightarrow R_3$.

c) Ningún número de intercambios de filas puede producir una matriz escalonada.

1.21. Reducir por filas la siguiente matriz a forma escalonada:

$$A = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{pmatrix}$$

Utilizamos $a_{11} = 1$ como pivote para obtener ceros bajo él; esto es, efectuamos las operaciones entre filas $-2R_1 + R_2 \to R_2$ y $-3R_1 + R_3 \to R_3$ para obtener la matriz

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 3 \end{pmatrix}$$

Ahora utilizamos $a_{23} = 4$ como pivote para obtener un cero bajo él; esto es, efectuamos la operación entre filas $-5R_2 + 4R_3 \to R_3$ para obtener la matriz

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

que está en forma escalonada.

1.22. Reducir por filas la siguiente matriz a forma escalonada:

$$B = \begin{pmatrix} -4 & 1 & -6 \\ 1 & 2 & -5 \\ 6 & 3 & -4 \end{pmatrix}$$

Los cálculos manuales suelen ser más simples si el elemento pivote es igual a 1. Por consiguiente, primero intercambiamos R_1 y R_2; luego efectuamos $4R_1 + R_2 \to R_2$ y $-6R_1 + R_3 \to R_3$; y entonces efectuamos $R_2 + R_3 \to R_3$:

$$B \sim \begin{pmatrix} 1 & 2 & -5 \\ -4 & 1 & -6 \\ 6 & 3 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & -9 & 26 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{pmatrix}$$

La matriz está ahora en forma escalonada.

1.23. Describir el algoritmo de reducción por filas *pivotando*. Además, describir las ventajas, si las hay, de utilizar este algoritmo.

El algoritmo de reducción por filas se convierte en un algoritmo de pivotar si la entrada de mayor valor absoluto en la columna j se elige como el pivote a_{1j_1} y si se utiliza la operación entre filas

$$(-a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$$

La principal ventaja del algoritmo de pivotar es que la operación entre filas anterior sólo implica división por el pivote a_{1j_1} y en las computadoras los errores de redondeo pueden ser sustancialmente reducidos cuando se divide por un número tan grande en valor absoluto como sea posible.

1.24. Usar el algoritmo de pivotar para reducir la siguiente matriz A a forma escalonada:

$$A = \begin{pmatrix} 2 & -2 & 2 & 1 \\ -3 & 6 & 0 & -1 \\ 1 & -7 & 10 & 2 \end{pmatrix}$$

Primero intercambiamos R_1 y R_2 de modo que -3 pueda ser utilizado como el pivote y entonces efectuamos $(\frac{2}{3})R_1 + R_2 \to R_2$ y $(\frac{1}{3})R_1 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} -3 & 6 & 0 & -1 \\ 2 & -2 & 2 & 1 \\ 1 & -7 & 10 & 2 \end{pmatrix} \sim \begin{pmatrix} -3 & 6 & 0 & -1 \\ 0 & 2 & 2 & \frac{1}{3} \\ 0 & -5 & 10 & \frac{5}{3} \end{pmatrix}$$

Ahora intercambiamos R_2 y R_3 de modo que -5 pueda ser utilizado como el pivote y efectuamos $(\frac{2}{5})R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 2 & 2 & \frac{1}{3} \end{pmatrix} \sim \begin{pmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 0 & 6 & 1 \end{pmatrix}$$

Se ha llevado la matriz a forma escalonada.

FORMA CANONICA POR FILAS

1.25. ¿Cuáles de las siguientes matrices están en forma canónica por filas?

$$\begin{pmatrix} 1 & 2 & -3 & 0 & 1 \\ 0 & 0 & 5 & 2 & -4 \\ 0 & 0 & 0 & 7 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 7 & -5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 5 & 0 & 2 \\ 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 7 \end{pmatrix}$$

La primera matriz no está en forma canónica por filas puesto que, por ejemplo, dos de las entradas principales no nulas son 5 y 7 en lugar de 1. Además, hay entradas distintas de cero sobre las entradas principales no nulas 5 y 7. Las matrices segunda y tercera están en forma canónica por filas.

1.26. Reducir la siguiente matriz a forma canónica por filas:

$$B = \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 6 & 6 & 0 & 20 & 19 \end{pmatrix}$$

Primero reducimos B a forma escalonada efectuando $-2R_1 + R_2 \to R_2$ y $-3R_1 + R_3 \to R_3$ y entonces $-R_2 + R_3 \to R_3$:

$$B \sim \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 3 & 2 & 7 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 4 & 2 \end{pmatrix}$$

A continuación reducimos la matriz escalonada a forma canónica por filas. Específicamente, multiplicamos primero R_3 por $\frac{1}{4}$, de modo que el pivote sea $b_{34} = 1$, y entonces efectuamos $2R_3 + R_2 \to R_2$ y $-6R_3 + R_1 \to R_1$:

$$B \sim \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

Ahora multiplicamos R_2 por $\frac{1}{3}$, haciendo el pivote $b_{23} = 1$, y efectuamos $R_2 + R_1 \to R_1$:

$$B \sim \begin{pmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

Finalmente, multiplicamos R_1 por $\frac{1}{2}$ para obtener la forma canónica por filas

$$B \sim \begin{pmatrix} 1 & 1 & 0 & 0 & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

1.27. Reducir la siguiente matriz a forma canónica por filas:

$$A = \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 1 & 1 & 4 & -1 & 3 \\ 2 & 5 & 9 & -2 & 8 \end{pmatrix}$$

Primero reducimos A a forma escalonada efectuando $-R_1 + R_2 \to R_2$ y $-2R_1 + R_3 \to R_3$ y entonces efectuando $-3R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 9 & 3 & -4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

Ahora utilizamos sustitución hacia atrás. Multiplicamos R_3 por $\frac{1}{2}$ para obtener el pivote $a_{34} = 1$ y entonces efectuamos $2R_3 + R_2 \to R_2$ y $-R_3 + R_1 \to R_1$:

$$A \sim \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & 0 & \frac{3}{2} \\ 0 & 3 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

Ahora multiplicamos R_2 por $\frac{1}{3}$ para obtener el pivote $a_{22} = 1$ y entonces efectuamos $2R_2 + R_1 \to R_1$:

$$A \sim \begin{pmatrix} 1 & -2 & 3 & 0 & \frac{3}{2} \\ 0 & 1 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \frac{11}{3} & 0 & \frac{17}{6} \\ 0 & 1 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

Como $a_{11} = 1$, la última matriz es la forma canónica por filas deseada.

1.28. Describir el algoritmo de eliminación de Gauss-Jordan, que reduce una matriz arbitraria A a su forma canónica por filas.

El algoritmo de Gauss-Jordan es similar al de eliminación gaussiana, salvo que aquí primero se normaliza una fila para obtener un pivote unidad y a continuación se utiliza el pivote para situar ceros tanto debajo como encima de él antes de obtener el pivote siguiente.

1.29. Utilizar la eliminación de Gauss-Jordan para obtener la forma canónica por filas de la matriz del Problema 1.27.

Usamos la entrada principal no nula $a_{11} = 1$ como pivote para poner ceros debajo, efectuando $-R_1 + R_2 \to R_2$ y $-2R_1 + R_3 \to R_3$; esto proporciona

$$A \sim \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 9 & 3 & -4 & 4 \end{pmatrix}$$

Multiplicamos R_2 por $\frac{1}{3}$ para conseguir el pivote $a_{22} = 1$ y obtenemos ceros encima y debajo de a_{22} efectuando $-9R_2 + R_3 \to R_3$ y $2R_2 + R_1 \to R_1$:

$$A \sim \begin{pmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 9 & 3 & -4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

Por último, multiplicamos R_3 por $\frac{1}{2}$ para conseguir el pivote $a_{34} = 1$ y obtenemos ceros sobre a_{34} efectuando $(\frac{2}{3})R_3 + R_2 \to R_2$ y $(\frac{1}{3})R_3 + R_1 \to R_1$:

$$A \sim \begin{pmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \frac{11}{3} & 0 & \frac{17}{6} \\ 0 & 1 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

1.30. Se habla de «una» forma escalonada de una matriz A y de «la» forma canónica por filas de A. ¿Por qué?

Una matriz arbitraria A puede ser equivalente por filas a muchas matrices escalonadas. Por el contrario, con independencia del algoritmo utilizado, una matriz A es equivalente por filas a una única matriz en forma canónica por filas. (El término «canónico» usualmente connota unicidad.) Por ejemplo, las formas canónicas por filas de los Problemas 1.27 y 1.29 son iguales.

1.31. Dada una matriz escalonada $n \times n$ en forma triangular,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

con todos los $a_{ii} \neq 0$, hallar la forma canónica por filas de A.

Multiplicando R_n por $1/a_{nn}$ y utilizando el nuevo $a_{nn} = 1$ como pivote obtenemos la matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & 0 \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & 0 \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Obsérvese que la última columna de A se ha convertido en un vector unitario. Cada sustitución hacia atrás subsiguiente proporciona un nuevo vector columna unitario y el resultado final es

$$A \sim \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

es decir, *A tiene la matriz identidad, I, $n \times n$,* como su forma canónica por filas.

1.32. Reducir la siguiente matriz triangular con elementos diagonales no nulos a forma canónica por filas:

$$C = \begin{pmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{pmatrix}$$

Por el Problema 1.31, C es equivalente por filas a la matriz identidad. De forma alternativa, por sustitución hacia atrás,

$$C \sim \begin{pmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 5 & -9 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 5 & -9 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

SISTEMAS DE ECUACIONES LINEALES EN FORMA MATRICIAL

1.33. Hallar la matriz ampliada M y la matriz de coeficientes A del siguiente sistema:

$$\begin{aligned} x + 2y - 3z &= 4 \\ 3y - 4z + 7x &= 5 \\ 6z + 8x - 9y &= 1 \end{aligned}$$

Empezaremos alineando las incógnitas del sistema para obtener

$$\begin{aligned} x + 2y - 3z &= 4 \\ 7x + 3y - 4z &= 5 \\ 8x - 9y + 6z &= 1 \end{aligned}$$

Entonces

$$M = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 7 & 3 & -4 & 5 \\ 8 & -9 & 6 & 1 \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} 1 & 2 & -3 \\ 7 & 3 & -4 \\ 8 & -9 & 6 \end{pmatrix}$$

1.34. Resolver, utilizando la matriz ampliada,

$$\begin{aligned} x - 2y + 4z &= 2 \\ 2x - 3y + 5z &= 3 \\ 3x - 4y + 6z &= 7 \end{aligned}$$

Reducimos la matriz ampliada a forma escalonada:

$$\begin{pmatrix} 1 & -2 & 4 & 2 \\ 2 & -3 & 5 & 3 \\ 3 & -4 & 6 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 2 & -6 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

La tercera fila de la matriz escalonada corresponde a la ecuación degenerada $0 = 3$; en consecuencia, el sistema no tiene solución.

1.35. Resolver, utilizando la matriz ampliada,

$$\begin{aligned} x + 2y - 3z - 2s + 4t &= 1 \\ 2x + 5y - 8z - s + 6t &= 4 \\ x + 4y - 7z + 5s + 2t &= 8 \end{aligned}$$

Reducimos la matriz ampliada a forma escalonada y luego a forma canónica por filas:

$$\begin{pmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 2 & 5 & -8 & -1 & 6 & 4 \\ 1 & 4 & -7 & 5 & 2 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 2 & -4 & 7 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & -3 & 0 & 8 & 7 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 0 & 24 & 21 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{pmatrix}$$

Así la solución en forma de variables libres es

$$\begin{aligned} x + z + 24t &= 21 \\ y - 2z - 8t &= -7 \\ s + 2t &= 3 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x &= 21 - z - 24t \\ y &= -7 + 2z + 8t \\ s &= 3 - 2t \end{aligned}$$

donde z y t son las variables libres.

1.36. Resolver, utilizando la matriz ampliada,

$$\begin{aligned} x + 2y - z &= 3 \\ x + 3y + z &= 5 \\ 3x + 8y + 4z &= 17 \end{aligned}$$

Reducimos la matriz ampliada a forma escalonada y luego a forma canónica por filas:

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 1 & 3 & 1 & 5 \\ 3 & 8 & 4 & 17 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 7 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 3 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & \frac{4}{3} \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 0 & \frac{13}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{pmatrix}$$

El sistema tiene la solución única $x = \frac{17}{3}$, $y = -\frac{2}{3}$, $z = \frac{4}{3}$ o $u = (\frac{17}{3}, -\frac{2}{3}, \frac{4}{3})$.

SISTEMAS HOMOGENEOS

1.37. Determinar si cada uno de los sistemas siguientes tiene una solución no nula.

$$\begin{aligned} x - 2y + 3z - 2w &= 0 \\ 3x - 7y - 2z + 4w &= 0 \\ 4x + 3y + 5z + 2w &= 0 \end{aligned}$$

a)

$$\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned}$$

b)

$$\begin{aligned} x + 2y - z &= 0 \\ 2x + 5y + 2z &= 0 \\ x + 4y + 7z &= 0 \\ x + 3y + 3z &= 0 \end{aligned}$$

c)

a) El sistema debe tener una solución no nula puesto que hay más incógnitas que ecuaciones.

b) Reducimos a forma escalonada:

$$\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned} \quad \text{a} \quad \begin{aligned} x + 2y - 3z &= 0 \\ y + 8z &= 0 \\ -7y + 5z &= 0 \end{aligned} \quad \text{a} \quad \begin{aligned} x + 2y - 3z &= 0 \\ y + 8z &= 0 \\ 61z &= 0 \end{aligned}$$

En forma escalonada hay exactamente tres ecuaciones con tres incógnitas; por tanto, el sistema tiene una solución única, la solución nula.

c) Reducimos a forma escalonada:

$$\begin{aligned} x + 2y - z &= 0 \\ 2x + 5y + 2z &= 0 \\ x + 4y + 7z &= 0 \\ x + 3y + 3z &= 0 \end{aligned} \quad \text{a} \quad \begin{aligned} x + 2y - z &= 0 \\ y + 4z &= 0 \\ 2y + 8z &= 0 \\ y + 4z &= 0 \end{aligned} \quad \text{a} \quad \begin{aligned} x + 2y - z &= 0 \\ y + 4z &= 0 \end{aligned}$$

En forma escalonada hay sólo dos ecuaciones con tres incógnitas; por consiguiente, el sistema tiene una solución no nula.

1.38. Encontrar la dimensión y una base para la solución general W del sistema homogéneo

$$\begin{aligned} x + 3y - 2z + 5s - 3t &= 0 \\ 2x + 7y - 3z + 7s - 5t &= 0 \\ 3x + 11y - 4z + 10s - 9t &= 0 \end{aligned}$$

Mostrar cómo la base da la forma paramétrica de la solución general del sistema.

Reducimos el sistema a forma escalonada. Efectuamos las operaciones $-2L_1 + L_2 \to L_2$ y $-3L_1 + L_3 \to L_3$ y entonces $-2L_2 + L_3 \to L_3$ para obtener

$$\begin{aligned} x + 3y - 2z + 5s - 3t &= 0 \\ y + z - 3s + t &= 0 \\ 2y + 2z - 5s &= 0 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x + 3y - 2z + 5s - 3t &= 0 \\ y + z - 3s + t &= 0 \\ s - 2t &= 0 \end{aligned}$$

En forma escalonada, el sistema tiene dos variables libres, z y t; por ende, $\dim W = 2$. Puede obtenerse una base $[u_1, u_2]$ para W como sigue:

1. Hacemos $z = 1$, $t = 0$. La sustitución hacia atrás proporciona $s = 0$, luego $y = -1$ y después $x = 5$. Por tanto, $u_1 = (5, -1, 1, 0, 0)$.

2. Hacemos $z = 0$, $t = 1$. La sustitución hacia atrás proporciona $s = 2$, luego $y = 5$ y después $x = -22$. Por consiguiente, $u_2 = (-22, 5, 0, 2, 1)$.

Multiplicando los vectores de la base por los parámetros a y b, respectivamente, obtenemos

$$au_1 + bu_2 = a(5, -1, 1, 0, 0) + b(-22, 5, 0, 2, 1) = (5a - 22b, -a + 5b, a, 2b, b)$$

Esta es la forma paramétrica de la solución general.

1.39. Hallar la dimensión y una base para la solución general W del sistema homogéneo

$$\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned}$$

Reducimos el sistema a forma escalonada. Del Problema 1.37 *b*) tenemos

$$\begin{aligned} x + 2y - 3z &= 0 \\ y + 8z &= 0 \\ 61z &= 0 \end{aligned}$$

No hay variables libres (el sistema está en forma triangular). Luego $\dim W = 0$ y W no tiene base. Específicamente, W consiste sólo en la solución nula, $W = \{0\}$.

1.40. Encontrar la dimensión y una base para la solución general W del sistema homogéneo

$$\begin{aligned} 2x + 4y - 5z + 3t &= 0 \\ 3x + 6y - 7z + 4t &= 0 \\ 5x + 10y - 11z + 6t &= 0 \end{aligned}$$

Reducimos el sistema a forma escalonada. Efectuamos $-3L_1 + 2L_2 \to L_2$ y $-5L_1 + 2L_3 \to L_3$ y entonces $-3L_2 + L_3 \to L_3$ para obtener

$$\begin{aligned} 2x + 4y - 5z + 3t &= 0 \\ z - t &= 0 \\ 3z - 3t &= 0 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 2x + 4y - 5z + 3t &= 0 \\ z - t &= 0 \end{aligned}$$

En forma escalonada, el sistema tiene dos variables libres, y y t; por consiguiente, $\dim W = 2$. Una base $\{u_1, u_2\}$ de W puede obtenerse como sigue:

1. Tomamos $y = 1$, $t = 0$. La sustitución hacia atrás proporciona la solución $u_1 = (-2, 1, 0, 0)$.
2. Tomamos $y = 0$, $t = 1$. La sustitución hacia atrás proporciona la solución $u_2 = (1, 0, 1, 1)$.

1.41. Considérese el sistema

$$\begin{aligned} x - 3y - 2z + 4t &= 5 \\ 3x - 8y - 3z + 8t &= 18 \\ 2x - 3y + 5z - 4t &= 19 \end{aligned}$$

a) Hallar la forma paramétrica de la solución general del sistema.

b) Mostrar que el resultado de *a*) puede reescribirse en la forma dada por el Teorema 1.11.

a) Reducimos el sistema a forma escalonada. Efectuamos $-3L_1 + L_2 \to L_2$ y $-2L_1 + L_3 \to L_3$ y entonces $-3L_2 + L_3 \to L_3$ para obtener

$$\begin{aligned} x - 3y - 2z + 4t &= 5 \\ y + 3z - 4t &= 3 \\ 3y + 9z - 12t &= 9 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x - 3y - 2z + 4t &= 5 \\ y + 3z - 4t &= 3 \end{aligned}$$

En forma escalonada, las variables libres son z y t. Tomamos $z = a$ y $t = b$, donde a y b son parámetros. La sustitución hacia atrás proporciona $y = 3 - 3a + 4b$ y entonces $x = 14 - 7a + 8b$. De este modo, la forma paramétrica de la solución es

$$x = 14 - 7a + 8b \qquad y = 3 - 3a + 4b \qquad z = a \qquad t = b \qquad (*)$$

b) Sea $v_0 = (14, 3, 0, 0)$ el vector de términos constantes en $(*)$; sea $u_1 = (-7, -3, 1, 0)$ el vector de coeficientes de a en $(*)$, y sea $u_2 = (8, 4, 0, 1)$ el vector de coeficientes de b en $(*)$. Entonces la solución general $(*)$ puede reescribirse en forma vectorial como

$$(x, y, z, t) = v_0 + au_1 + bu_2 \qquad (**)$$

A continuación probamos que $(**)$ es la solución general del Teorema 1.11. En primer lugar, nótese que v_0 es la solución del sistema inhomogéneo obtenido haciendo $a = 0$ y $b = 0$. Consideremos el sistema homogéneo asociado, en forma escalonada:

$$\begin{aligned} x - 3y - 2z + 4t &= 0 \\ y + 3z - 4t &= 0 \end{aligned}$$

Las variables libres son z y t. Hacemos $z = 1$ y $t = 0$ para obtener la solución $u_1 = (-7, -3, 1, 0)$. Hacemos ahora $z = 0$ y $t = 1$ para obtener la solución $u_2 = (8, 4, 0, 1)$. Por el Teorema 1.10, $\{u_1, u_2\}$ es una base del espacio solución del sistema homogéneo asociado. Entonces $(**)$ tiene la forma deseada.

PROBLEMAS VARIOS

1.42. Demostrar que cada una de las operaciones elementales $[E_1]$, $[E_2]$, $[E_3]$ tiene una operación inversa del mismo tipo.

$[E_1]$ Intercambiar las ecuaciones i-ésima y j-ésima: $L_i \leftrightarrow L_j$.

$[E_2]$ Multiplicar la ecuación i-ésima por un escalar no nulo k: $kL_i \to L_i$, $k \neq 0$.

$[E_3]$ Sustituir la ecuación i-ésima por ella misma más k veces la j-ésima: $kL_j + L_i \to L_i$.

a) Intercambiando el mismo par de ecuaciones dos veces obtenemos el sistema original. Esto es, $L_i \leftrightarrow L_j$ es su propia inversa.

b) Multiplicando la i-ésima ecuación por k y luego por k^{-1}, o por k^{-1} y luego por k, obtenemos el sistema original. En otras palabras, las operaciones $kL_i \to L_i$ y $k^{-1}L_i \to L_i$ son inversas.

c) Efectuando la operación $kL_j + L_i \to L_i$ y luego la $-kL_j + L_i \to L_i$, o viceversa, obtenemos el sistema original. En otras palabras, las operaciones $kL_j + L_i \to L_i$ y $-kL_j + L_i \to L_i$ son inversas.

1.43. Demostrar que el efecto de ejecutar la siguiente operación $[E]$ puede obtenerse efectuando $[E_2]$ y luego $[E_3]$.

$[E]$ Sustituir la ecuación i-ésima por k (no nulo) veces ella misma más k' veces la j-ésima: $k'L_j + kL_i \to L_i$, $k \neq 0$.

Efectuar $kL_i \to L_i$ y luego $k'L_j + L_i \to L_i$ conduce al mismo resultado que efectuar la operación $k'L_j + kL_i \to L_i$.

1.44. Supongamos que cada ecuación L_i en el sistema [1.3] se multiplica por una constante c_i y que las ecuaciones resultantes se suman para dar

$$(c_1 a_{11} + \cdots + c_m a_{m1})x_1 + \cdots + (c_1 a_{1n} + \cdots + c_m a_{mn})x_n = c_1 b_1 + \cdots + c_m b_m \qquad [1]$$

Tal ecuación se denomina una combinación lineal de las ecuaciones L_i. Demostrar que cualquier solución del sistema [1.3] es también solución de la combinación lineal [1].

Supongamos que $u = (k_1, k_2, \ldots, k_n)$ es una solución de [1.3]:

$$a_{i1}k_1 + a_{i2}k_2 + \cdots + a_{in}k_n = b_i \qquad (i = 1, \ldots, m) \qquad [2]$$

Para probar que u es una solución de [1] debemos verificar la ecuación

$$(c_1 a_{11} + \cdots + c_m a_{m1})k_1 + \cdots + (c_1 a_{1n} + \cdots + c_m a_{mn})k_n = c_1 b_1 + \cdots + c_m b_m$$

Pero ésta puede arreglarse de la forma

$$c_1(a_{11}k_1 + \cdots + a_{1n}k_n) + \cdots + c_m(a_{m1} + \cdots + a_{mn}k_n) = c_1 b_1 + \cdots + c_m b_m$$

o por [2]

$$c_1 b_1 + \cdots + c_m b_m = c_1 b_1 + \cdots + c_m b_m$$

que es claramente una expresión cierta.

1.45. Supóngase que un sistema de ecuaciones lineales (#) se obtiene a partir de otro (*) mediante la ejecución de una sola operación elemental $-[E_1]$, $[E_2]$, o $[E_3]-$. Demostrar que (#) y (*) tienen todas las soluciones en común.

Cada ecuación de (#) es una combinación lineal de ecuaciones de (*). Por tanto, por el Problema. 1.44, cualquier solución de (*) será solución de todas las ecuaciones de (#). Dicho de otro modo, el conjunto solución de (*) está contenido en el de (#). Por otra parte, como las operaciones $[E_1]$, $[E_2]$ y $[E_3]$ tienen operaciones elementales inversas, el sistema (*) puede obtenerse del (#) por medio de una sola operación elemental. De acuerdo con ello, el conjunto solución de (#) está contenido en el de (*). Siendo así, (#) y (*) tienen las mismas soluciones.

1.46. Demostrar el Teorema 1.4.

Por el Problema 1.45, cada paso deja inalterado el conjunto solución. Por tanto, el sistema original (#) y el final (*) tienen el mismo conjunto solución.

1.47. Probar que las tres afirmaciones siguientes, relativas a un sistema de ecuaciones lineales, son equivalentes: i) El sistema es compatible (tiene solución). ii) Ninguna combinación lineal de las ecuaciones es la ecuación

$$0x_1 + 0x_2 + \cdots + 0x_n = b \neq 0 \qquad (*)$$

iii) El sistema es reducible a forma escalonada.

Si el sistema es reducible a forma escalonada, dicha forma tiene una solución y por ende la tiene el sistema original. Así iii) implica i).

Si el sistema tiene una solución, por el Problema 1.44 cualquier combinación lineal de las ecuaciones tiene también una solución. Pero (*) no tiene solución, luego no es combinación lineal de las ecuaciones. Así i) implica ii).

Finalmente, supongamos que el sistema no es reducible a forma escalonada. Entonces en el algoritmo gaussiano debe dar una ecuación de la forma (*). Por consiguiente, (*) es una combinación lineal de las ecuaciones. Así no- iii) implica no- ii) o, equivalentemente, ii) implica iii).

PROBLEMAS SUPLEMENTARIOS

RESOLUCION DE ECUACIONES LINEALES

1.48. Resolver:

a) $\begin{aligned} 2x + 3y &= 1 \\ 5x + 7y &= 3 \end{aligned}$ *b)* $\begin{aligned} 2x + 4y &= 10 \\ 3x + 6y &= 15 \end{aligned}$ *c)* $\begin{aligned} 4x - 2y &= 5 \\ -6x + 3y &= 1 \end{aligned}$

1.49. Resolver:

a) $\begin{aligned} 2x - y - 3z &= 5 \\ 3x - 2y + 2z &= 5 \\ 5x - 3y - z &= 16 \end{aligned}$ *b)* $\begin{aligned} 2x + 3y - 2z &= 5 \\ x - 2y + 3z &= 2 \\ 4x - y + 4z &= 1 \end{aligned}$ *c)* $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 3x + 2y + 17z &= 1 \end{aligned}$

1.50. Resolver:

a) $\begin{aligned} 2x + 3y &= 3 \\ x - 2y &= 5 \\ 3x + 2y &= 7 \end{aligned}$ *b)* $\begin{aligned} x + 2y - 3z + 2t &= 2 \\ 2x + 5y - 8z + 6t &= 5 \\ 3x + 4y - 5z + 2t &= 4 \end{aligned}$ *c)* $\begin{aligned} x + 2y - z + 3t &= 3 \\ 2x + 4y + 4z + 3t &= 9 \\ 3x + 6y - z + 8t &= 10 \end{aligned}$

1.51. Resolver:

$$a)\quad \begin{aligned} x + 2y + 2z &= 2 \\ 3x - 2y - z &= 5 \\ 2x - 5y + 3z &= -4 \\ x + 4y + 6z &= 0 \end{aligned} \qquad b)\quad \begin{aligned} x + 5y + 4z - 13t &= 3 \\ 3x - y + 2z + 5t &= 2 \\ 2x + 2y + 3z - 4t &= 1 \end{aligned}$$

SISTEMAS HOMOGENEOS

1.52. Determinar si cada uno de los siguientes sistemas tiene una solución no nula:

$$a)\quad \begin{aligned} x + 3y - 2z &= 0 \\ x - 8y + 8z &= 0 \\ 3x - 2y + 4z &= 0 \end{aligned} \qquad b)\quad \begin{aligned} x + 3y - 2z &= 0 \\ 2x - 3y + z &= 0 \\ 3x - 2y + 2z &= 0 \end{aligned} \qquad c)\quad \begin{aligned} x + 2y - 5z + 4t &= 0 \\ 2x - 3y + 2z + 3t &= 0 \\ 4x - 7y + z - 6t &= 0 \end{aligned}$$

1.53. Hallar la dimensión y una base para la solución general W de cada uno de los sistemas siguientes.

$$a)\quad \begin{aligned} x + 3y + 2z - s - t &= 0 \\ 2x + 6y + 5z + s - t &= 0 \\ 5x + 15y + 12z + s - 3t &= 0 \end{aligned} \qquad b)\quad \begin{aligned} 2x - 4y + 3z - s + 2t &= 0 \\ 3x - 6y + 5z - 2s + 4t &= 0 \\ 5x - 10y + 7z - 3s + t &= 0 \end{aligned}$$

MATRICES ESCALONADAS Y OPERACIONES ELEMENTALES ENTRE FILAS

1.54. Reducir A a forma escalonada y luego a su forma canónica por filas, siendo

$$a)\quad A = \begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{pmatrix} \qquad b)\quad A = \begin{pmatrix} 2 & 3 & -2 & 5 & 1 \\ 3 & -1 & 2 & 0 & 4 \\ 4 & -5 & 6 & -5 & 7 \end{pmatrix}$$

1.55. Reducir A a forma escalonada y luego a su forma canónica por filas, siendo

$$a)\quad A = \begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{pmatrix} \qquad b)\quad A = \begin{pmatrix} 0 & 1 & 3 & -2 \\ 0 & 4 & -1 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 5 & -3 & 4 \end{pmatrix}$$

1.56. Describir todas las matrices 2×2 posibles que estén en forma escalonada reducida por filas.

1.57. Supóngase que A es una matriz cuadrada escalonada reducida por filas. Demostrar que si $A \neq I$, la matriz identidad, A tiene una fila nula.

1.58. Demostrar que cada una de las operaciones elementales entre filas siguientes tiene una operación inversa del mismo tipo.

[E_1] Intercambiar las filas i-ésima y j-ésima: $R_i \leftrightarrow R_j$.

[E_2] Multiplicar la fila i-ésima por un escalar no nulo k: $kR_i \to R_i$, $k \neq 0$.

[E_3] Sustituir la fila i-ésima por ella misma más k veces la j-ésima: $kR_j + R_i \to R_i$.

1.59. Demostrar que la equivalencia por filas es una relación de equivalencia:

i) A es equivalente por filas a A.

ii) Si A es equivalente por filas a B, entonces B es equivalente por filas a A.

iii) Si A es equivalente por filas a B y B es equivalente por filas a C, entonces A es equivalente por filas a C.

PROBLEMAS VARIOS

1.60. Considérense dos ecuaciones lineales generales con dos incógnitas x e y sobre el cuerpo real **R**:

$$ax + by = e$$
$$cx + dy = f$$

Demostrar que:

i) si $\frac{a}{c} \neq \frac{b}{d}$, es decir, si $ad - bc \neq 0$, entonces el sistema tiene la solución única $x = \frac{de - bf}{ad - bc}$, $y = \frac{af - ce}{ad - bc}$;

ii) si $\frac{a}{c} = \frac{b}{d} \neq \frac{e}{f}$, entonces el sistema no tiene solución;

iii) si $\frac{a}{c} = \frac{b}{d} = \frac{e}{f}$, entonces el sistema tiene más de una solución.

1.61. Considérese el sistema

$$ax + by = 1$$
$$cx + dy = 0$$

Demostrar que si $ad - bc \neq 0$, entonces el sistema tiene la solución única $x = d/(ad - bc)$, $y = -c/(ad - bc)$. Demostrar también que si $ad - bc = 0$, $c \neq 0$ o $d \neq 0$, entonces el sistema no tiene solución.

1.62. Demostrar que una ecuación de la forma $0x_1 + 0x_2 + \cdots + 0x_n = 0$ puede añadirse a o suprimirse de un sistema sin alterar el conjunto solución.

1.63. Considérese un sistema de ecuaciones lineales con el mismo número de ecuaciones que de incógnitas:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array} \qquad [1]$$

i) Suponiendo que el sistema homogéneo asociado tiene sólo la solución nula, probar que [1] tiene una solución única para cada elección de las constantes b_i.

ii) Suponiendo que el sistema homogéneo asociado tiene una solución no nula, probar que existen constantes b_i para las que [1] no tiene solución. Demostrar también que si [1] tiene una solución, entonces tiene más de una.

1.64. Suponiendo que en un sistema de ecuaciones lineales homogéneo los coeficientes de una de las incógnitas son todos nulos, demostrar que el sistema tiene solución no nula.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

1.48. *a*) $x = 2,\ y = -1$; *b*) $x = 5 - 2a,\ y = a$; *c*) sin solución.

1.49. *a*) $(1, -3, -2)$; *b*) sin solución; *c*) $(-1 - 7a, 2 + 2a, a)$ o $\begin{cases} x = -1 - 7z \\ y = 2 + 2z \end{cases}$

1.50. *a*) $x = 3,\ y = -1$; *b*) $(-a + 2b, 1 + 2a - 2b, a, b)$ o $\begin{cases} x = -z + 2t \\ y = 1 + 2z - 2t \end{cases}$

c) $(\frac{7}{2} - 5b/2 - 2a, a, \frac{1}{2} + b/2, b)$ o $\begin{cases} x = \frac{7}{2} - 5t/2 - 2y \\ z = \frac{1}{2} + t/2 \end{cases}$

1.51. *a*) $(2, 1, -1)$; *b*) sin solución.

1.52. *a*) sí; *b*) no; *c*) sí, por el Teorema 1.8.

1.53. *a*) $\dim W = 3$; $u_1 = (-3, 1, 0, 0, 0)$, $u_2 = (7, 0, -3, 1, 0)$, $u_3 = (3, 0, -1, 0, 1)$.
b) $\dim W = 2$; $u_1 = (-2, 1, 0, 0, 0)$, $u_2 = (5, 0, -5, -3, 1)$.

1.54. *a*) $\begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 0 & -6 & 1 \end{pmatrix}$ y $\begin{pmatrix} 1 & 2 & 0 & 0 & \frac{4}{3} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{6} \end{pmatrix}$;

b) $\begin{pmatrix} 2 & 3 & -2 & 5 & 1 \\ 0 & -11 & 10 & -15 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} 1 & 0 & \frac{4}{11} & \frac{5}{11} & \frac{13}{11} \\ 0 & 1 & -\frac{10}{11} & \frac{15}{11} & -\frac{5}{11} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

1.55. *a*) $\begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} 1 & 0 & \frac{4}{11} & \frac{13}{11} \\ 0 & 1 & -\frac{5}{11} & \frac{3}{11} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$;

b) $\begin{pmatrix} 0 & 1 & 3 & -2 \\ 0 & 0 & -13 & 11 \\ 0 & 0 & 0 & 35 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

CAPITULO 2

Vectores en $\mathbf{R}^n$ y $\mathbf{C}^n$. Vectores espaciales

2.1. INTRODUCCION

En varias aplicaciones físicas aparecen ciertas cantidades, tales como temperatura y rapidez (módulo de la velocidad), que poseen sólo «magnitud». Estas pueden representarse por números reales y se denominan *escalares*. Por otra parte, también existen cantidades, tales como fuerza y velocidad, que poseen tanto «magnitud» como «dirección». Estas pueden representarse por flechas (que tienen longitudes y direcciones apropiadas y parten de algún punto de referencia dado, O) y se denominan *vectores*.

Comenzamos considerando las siguientes operaciones con vectores.

i) ***Suma:*** La resultante $\mathbf{u} + \mathbf{v}$ de dos vectores $\mathbf{u}$ y $\mathbf{v}$ se obtiene por la llamada ley del paralelogramo; esto es, $\mathbf{u} + \mathbf{v}$ es la diagonal del paralelogramo determinado por $\mathbf{u}$ y $\mathbf{v}$, como se muestra en la Figura 2-1(a).

ii) ***Producto por un escalar:*** El producto $k\mathbf{u}$ de un número real k por un vector $\mathbf{u}$ se obtiene multiplicando la magnitud de $\mathbf{u}$ por k y manteniendo la misma dirección si $k \geqslant 0$, o la dirección opuesta si $k < 0$, tal y como se muestra en la Figura 2-1(b).

Suponemos que el lector está familiarizado con la representación de puntos en el plano mediante pares ordenados de números reales. Si se elige el origen de los ejes como el punto de referencia O mencionado más arriba, todo vector queda unívocamente determinado por las coordenadas de su extremo. La relación entre las operaciones anteriores y los extremos es la siguiente.

i) ***Suma:*** Si (a, b) y (c, d) son los extremos de los vectores $\mathbf{u}$ y $\mathbf{v}$, entonces $(a + c, b + d)$ será el extremo de $\mathbf{u} + \mathbf{v}$, como se muestra en la Figura 2-2(a).

ii) ***Producto por un escalar:*** Si (a, b) es el extremo del vector $\mathbf{u}$, entonces (ka, kb) será el del vector $k\mathbf{u}$, como se muestra en la Figura 2-2(b).

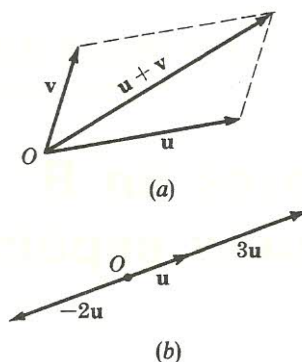


Figura 2-1.

Matemáticamente, identificamos el vector **u** con su extremo (a, b) y escribimos $\mathbf{u} = (a, b)$. Además llamamos al par ordenado de números reales (a, b) punto o vector, dependiendo de su interpretación. Generalizando esta noción, llamaremos vector a una n-pla $(a_1, a_2, \ldots, a_n)$ de números reales. En cualquier caso, puede utilizarse una notación particular para los vectores espaciales en $\mathbf{R}^3$ (Sección 2.8).

Presumimos al lector familiarizado con las propiedades elementales del cuerpo de los números reales que denotamos por **R**.

2.2. VECTORES EN $\mathbf{R}^n$

El conjunto de todas las n-plas de números reales, denotado por $\mathbf{R}^n$, recibe el nombre de *n-espacio*. Una n-pla particular en $\mathbf{R}^n$, digamos

$$u = (u_1, u_2, \ldots, u_n)$$

se denomina *punto* o *vector*; los números reales u_i son las *componentes* (o *coordenadas*) del vector u. Además, al discutir el espacio $\mathbf{R}^n$, utilizamos el término *escalar* para designar a los elementos de **R**.

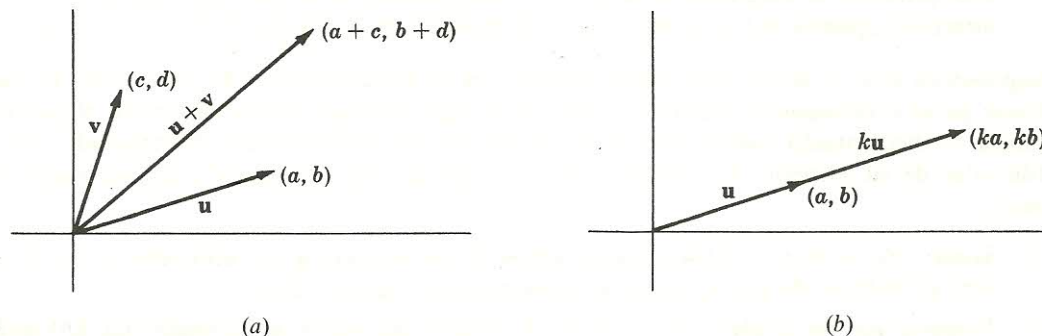


Figura 2-2.

Dos vectores u y v son *iguales*, escrito $u = v$, si tienen el mismo número de componentes, es decir, pertenecen al mismo espacio, y si las componentes correspondientes son iguales. Los vectores (1, 2, 3) y (2, 3, 1) no son iguales, puesto que no lo son las componentes correspondientes.

EJEMPLO 2.1

a) Considérense los siguientes vectores

$$(0, 1) \qquad (1, -3) \qquad (1, 2, \sqrt{3}, 4) \qquad (-5, \tfrac{1}{2}, 0, \pi)$$

Los dos primeros tienen dos componentes y por tanto son puntos en $\mathbf{R}^2$; los dos últimos tienen cuatro componentes y por tanto son puntos en $\mathbf{R}^4$.

b) Supóngase $(x - y, x + y, z - 1) = (4, 2, 3)$. Entonces, por definición de igualdad de vectores,

$$x - y = 4$$
$$x + y = 2$$
$$z - 1 = 3$$

Resolviendo el sistema de ecuaciones anterior se obtiene $x = 3$, $y = -1$ y $z = 4$.

A veces los vectores de un n-espacio se escriben como columnas en lugar de como se hizo anteriormente, como filas. Tales vectores se denominan *vectores columna*. Por ejemplo,

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ -3 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 7 \\ -8 \end{pmatrix} \qquad \begin{pmatrix} 1{,}2 \\ -35 \\ 28 \end{pmatrix}$$

son vectores columna con 2, 2, 3 y 3 componentes, respectivamente.

2.3. SUMA DE VECTORES Y PRODUCTO POR UN ESCALAR

Sean u y v vectores en $\mathbf{R}^n$:

$$u = (u_1, u_2, \ldots, u_n) \qquad \text{y} \qquad v = (v_1, v_2, \ldots, v_n)$$

La *suma* de u y v, escrito $u + v$, es el vector obtenido sumando las componentes correspondientes de éstos:

$$u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

El *producto* de un número real k por el vector u, escrito ku, es el vector obtenido multiplicando cada componente de u por k:

$$ku = (ku_1, ku_2, \ldots, ku_n)$$

Obsérvese que $u + v$ y ku son también vectores en $\mathbf{R}^n$. Definimos, además,

$$-u = -1u \qquad \text{y} \qquad u - v = u + (-v)$$

La suma de vectores con diferente número de componentes no está definida.

Las propiedades básicas de los vectores de $\mathbf{R}^n$ bajo las operaciones de suma vectorial y producto por un escalar se enuncian en el siguiente teorema (demostrado en el Problema 2.4). En el teorema, $0 \equiv (0, 0, \ldots, 0)$, el *vector nulo* de $\mathbf{R}^n$ (o *vector cero*).

Teorema 2.1: Para vectores arbitrarios u, v, $w \in \mathbf{R}^n$ y escalares arbitrarios k, $k' \in \mathbf{R}$,

i) $(u + v) + w = u + (v + w)$

ii) $u + 0 = u$

iii) $u + (-u) = 0$

iv) $u + v = v + u$

v) $k(u + v) = ku + kv$

vi) $(k + k')u = ku + k'u$

vii) $(kk')u = k(k'u)$

viii) $1u = u$

Supongamos que u y v son vectores en $\mathbf{R}^n$ y que $u = kv$ para algún escalar no nulo $k \in \mathbf{R}$. Entonces u se llama un *múltiplo* de v. Se dice que u está en la *misma dirección* que v si $k > 0$, y en la *dirección opuesta si* $k < 0$.

2.4. VECTORES Y ECUACIONES LINEALES

Dos importantes conceptos relativos a vectores, las combinaciones lineales y la dependencia lineal, están estrechamente relacionados con los sistemas de ecuaciones lineales, como se verá a continuación.

COMBINACIONES LINEALES

Consideremos un sistema inhomogéneo de m ecuaciones con n incógnitas:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}$$

Este sistema es equivalente a la siguiente ecuación vectorial:

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

esto es, la ecuación vectorial

$$x_1u_1 + x_2u_2 + \cdots + x_nu_n = v$$

donde $u_1, u_2, \ldots, u_n, v$ son los vectores columna anteriores, respectivamente.

Si el sistema en cuestión tiene una solución, se dice que v es una combinación lineal de los vectores u_i. Establezcamos formalmente este importante concepto.

Definición: Un vector v es una *combinación lineal* de vectores $u_1, u_2, \ldots, u_n$ si existen escalares $k_1, k_2, \ldots, k_n$ tales que

$$v = k_1u_1 + k_2u_2 + \cdots + k_nu_n$$

esto es, si la ecuación vectorial

$$v = x_1u_1 + x_2u_2 + \cdots + x_nu_n$$

tiene una solución cuando los x_i son escalares por determinar.

La definición anterior se aplica tanto a vectores columna como a vectores fila, aunque se haya ilustrado en términos de vectores columna.

EJEMPLO 2.2. Sean

$$v = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}, \qquad u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{y} \qquad u_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Entonces v es una combinación lineal de u_1, u_2 y u_3 ya que la ecuación vectorial (o sistema)

$$\begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{o} \qquad \begin{cases} 2 = x + y + z \\ 3 = x + y \\ -4 = x \end{cases}$$

tiene una solución $x = -4$, $y = 7$, $z = -1$. En otras palabras,

$$v = -4u_1 + 7u_2 - u_3$$

DEPENDENCIA LINEAL

Consideremos un sistema homogéneo de m ecuaciones con n incógnitas:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}$$

Este sistema es equivalente a la siguiente ecuación vectorial:

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

esto es,

$$x_1u_1 + x_2u_2 + \cdots + x_nu_n = 0$$

donde $u_1, u_2, \ldots, u_n$ son los vectores columna anteriores, respectivamente.

Si el sistema homogéneo anterior tiene una solución no nula, se dice que los vectores $u_1, u_2, \ldots, u_n$ son linealmente dependientes. Por el contrario, si el sistema tiene sólo la solución nula, se dice que los vectores son linealmente independientes. Enunciemos formalmente este importante concepto.

Definición: Los vectores $u_1, u_2, \ldots, u_n$ en $\mathbf{R}^n$ son *linealmente dependientes* si existen escalares $k_1, k_2, \ldots, k_n$, no todos nulos, tales que

$$k_1u_1 + k_2u_2 + \cdots + k_nu_n = 0$$

esto es, si la ecuación vectorial

$$x_1u_1 + x_2u_2 + \cdots + x_nu_n = 0$$

tiene una solución no nula donde los x_i son escalares por determinar. En caso contrario, se dice que los vectores son *linealmente independientes.*

La definición anterior se aplica tanto a vectores fila como a vectores columna, aunque se haya ilustrado en términos de vectores columna.

EJEMPLO 2.3

a) La única solución de

$$x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}1\\1\\0\end{pmatrix} + z\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \quad \text{o} \quad \begin{cases} x + y + z = 0 \\ x + y = 0 \\ x = 0 \end{cases}$$

es la solución nula $x = 0$, $y = 0$, $z = 0$. De aquí los tres vectores son linealmente independientes.

b) La ecuación vectorial (o sistema de ecuaciones lineales)

$$x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}2\\-1\\3\end{pmatrix} + z\begin{pmatrix}1\\-5\\3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \quad \text{o} \quad \begin{cases} x + 2y + z = 0 \\ x - y - 5z = 0 \\ x + 3y + 3z = 0 \end{cases}$$

tiene una solución no nula $(3, -2, 1)$, esto es, $x = 3$, $y = -2$, $z = 1$. De este modo, los tres vectores son linealmente dependientes.

2.5. PRODUCTO ESCALAR

Sean u y v vectores de $\mathbf{R}^n$:

$$u = (u_1, u_2, \ldots, u_n) \qquad \text{y} \qquad v = (v_1, v_2, \ldots, v_n)$$

El *producto escalar* o *interno* de u y v, denotado por $u \cdot v$, es el escalar obtenido multiplicando las componentes correspondientes de los vectores y sumando los productos resultantes:

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Se dice que los vectores u y v son *ortogonales* (o *perpendiculares*) si su producto escalar es cero, esto es, si $u \cdot v = 0$.

EJEMPLO 2.4. Sean $u = (1, -2, 3, -4)$, $v = (6, 7, 1, -2)$ y $w = (5, -4, 5, 7)$. Entonces

$$u \cdot v = 1 \cdot 6 + (-2) \cdot 7 + 3 \cdot 1 + (-4) \cdot (-2) = 6 - 14 + 3 + 8 = 3$$

$$u \cdot w = 1 \cdot 5 + (-2) \cdot (-4) + 3 \cdot 5 + (-4) \cdot 7 = 5 + 8 + 15 - 28 = 0$$

Entonces u y w son ortogonales.

Las propiedades básicas del producto escalar en $\mathbf{R}^n$ (demostradas en el Problema 2.17) son las siguientes.

Teorema 2.2: Para vectores arbitrarios $u, v, w \in \mathbf{R}^n$ y cualquier escalar $k \in \mathbf{R}$,

i) $(u + v) \cdot w = u \cdot w + v \cdot w$

ii) $(ku) \cdot v = k(u \cdot v)$

iii) $u \cdot v = v \cdot u$

iv) $u \cdot u \geqslant 0$, y $u \cdot u = 0$ si y sólo si $u = 0$

Nota: El espacio $\mathbf{R}^n$, con las operaciones anteriores de suma vectorial, producto por un escalar y producto interno, se suele llamar *n-espacio euclídeo*.

2.6. NORMA DE UN VECTOR

Sea $u = (u_1, u_2, \ldots, u_n)$ un vector en $\mathbf{R}^n$. La *norma* (o *longitud*) del vector u, escrito $\|u\|$, se define como la raíz cuadrada no negativa de $u \cdot u$:

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Como $u \cdot u \geqslant 0$, la raíz cuadrada existe. Además, si $u \neq 0$, entonces $\|u\| > 0$; y $\|0\| = 0$.

La definición anterior de norma de un vector se ajusta a la de longitud de un vector (flecha) en geometría (euclídea). Concretamente, supongamos que $\mathbf{u}$ es un vector (flecha) en el plano $\mathbf{R}^2$ con extremo $P(a, b)$, como se indica en la Figura 2-3. Entonces $|a|$ y $|b|$ son las longitudes de los lados del triángulo rectángulo determinado por $\mathbf{u}$ y las direcciones horizontal y vertical. Por el Teorema de Pitágoras, la longitud $|\mathbf{u}|$ de $\mathbf{u}$ es

$$|\mathbf{u}| = \sqrt{a^2 + b^2}$$

Este valor es igual a la norma de u anteriormente definida.

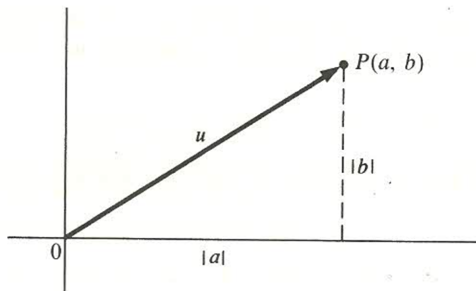


Figura 2-3.

EJEMPLO 2.5. Supongamos $u = (3, \; -12, \; -4)$. Para hallar $\|u\|$, calculamos primero $\|u\|^2 = u \cdot u$ elevando al cuadrado las componentes de u y sumándolas:

$$\|u\|^2 = 3^2 + (-12)^2 + (-4)^2 = 9 + 144 + 16 = 169$$

Entonces $\|u\| = \sqrt{169} = 13$.

Un vector u es unitario si $\|u\| = 1$ o, equivalentemente, si $u \cdot u = 1$. Ahora bien, si v es cualquier vector no nulo, entonces

$$\hat{v} \equiv \frac{1}{\|v\|} v = \frac{v}{\|v\|}$$

es un vector unitario en la misma dirección que v. (El proceso de hallar $\hat{v}$ se llama *normalizar* v.) Por ejemplo,

$$\hat{v} = \frac{v}{\|v\|} = \left(\frac{2}{\sqrt{102}}, \frac{-3}{\sqrt{102}}, \frac{8}{\sqrt{102}}, \frac{-5}{\sqrt{102}}\right)$$

es el vector unitario en la dirección del vector $v = (2, \; -3, \; 8, \; -5)$.

Establecemos ahora una relación fundamental (demostrada en el Problema 2.22) conocida como desigualdad de Cauchy-Schwarz.

Teorema 2.3 (Cauchy-Schwarz): Para vectores cualesquiera u, v en $\mathbf{R}^n$, $|u \cdot v| \leqslant \|u\| \; \|v\|$.

Utilizando la desigualdad anterior probaremos (Problema 2.33) el siguiente resultado, conocido como desigualdad triangular o de Minkowski.

Teorema 2.4 (Minkowski): Para vectores arbitrarios u, v en $\mathbf{R}^n$, $\|u + v\| \leqslant \|u\| + \|v\|$.

DISTANCIA. ANGULOS. PROYECCIONES

Sean $u = (u_1, u_2, \ldots, u_n)$ y $v = (v_1, v_2, \ldots, v_n)$ vectores en $\mathbf{R}^n$. La *distancia* entre u y v, denotada por $d(u, v)$, se define como

$$d(u, v) \equiv \|u - v\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

Probamos que esta definición corresponde a la noción usual de distancia euclídea en el plano $\mathbf{R}^2$. Consideremos $u = (a, b)$ y $v = (c, d)$ en $\mathbf{R}^2$. Como se muestra en la Figura 2-4, la distancia entre los puntos $P(a, b)$ y $Q(c, d)$ es

$$d = \sqrt{(a - c)^2 + (b - d)^2}$$

Por otra parte, utilizando la definición anterior,

$$d(u, v) = \|u - v\| = \|(a - c, b - d)\| = \sqrt{(a - c)^2 + (b - d)^2}$$

Ambas definiciones conducen al mismo valor.

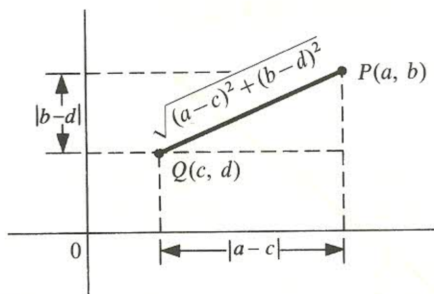


Figura 2-4.

Usando la desigualdad de Cauchy-Schwarz podemos definir el ángulo θ entre dos vectores no nulos u, v en $\mathbf{R}^n$ según

$$\cos \theta = \frac{u \cdot v}{\|u\| \|v\|}$$

Nótese que si $u \cdot v = 0$, entonces $\theta = 90°$ (o $\theta = \pi/2$). Esto está de acuerdo con nuestra definición previa de ortogonalidad.

EJEMPLO 2.6. Supongamos $u = (1, -2, 3)$ y $v = (3, -5, -7)$. Entonces

$$d(u, v) = \sqrt{(1 - 3)^2 + (-2 + 5)^2 + (3 + 7)^2} = \sqrt{4 + 9 + 100} = \sqrt{113}$$

Para hallar $\cos \theta$, donde θ es el ángulo entre u y v, calculamos primero

$$u \cdot v = 3 + 10 - 21 = -8 \qquad \|u\|^2 = 1 + 4 + 9 = 14 \qquad \|v\|^2 = 9 + 25 + 49 = 83$$

En tal caso

$$\cos \theta = \frac{u \cdot v}{\|u\| \|v\|} = -\frac{-8}{\sqrt{14}\sqrt{83}}$$

Sean u y $v \neq 0$ vectores en $\mathbf{R}^n$. La *proyección (vectorial) de u sobre v* es el vector

$$\text{proy}\,(u, v) = \frac{u \cdot v}{\|v\|^2}\, v$$

Probamos ahora que esta definición se ajusta a la noción de proyección vectorial utilizada en física. Consideremos los vectores **u** y **v** de la Figura 2-5. La proyección (perpendicular) de **u** sobre **v** es el vector **u***, de magnitud

$$|\mathbf{u}^*| = |\mathbf{u}| \cos \theta = |\mathbf{u}| \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|}$$

Para obtener **u*** multiplicamos su magnitud por el vector unitario en la dirección **v**:

$$\mathbf{u}^* = |\mathbf{u}^*| \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2} \mathbf{v}$$

Esto concuerda con la definición anterior de proy (u, v).

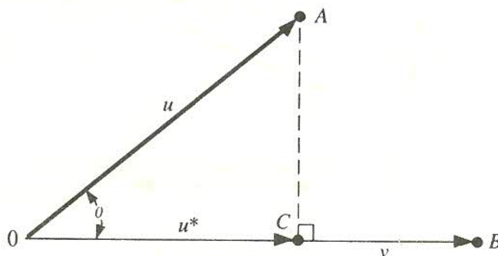


Figura 2-5.

EJEMPLO 2.7. Supongamos $u = (1, -2, 3)$ y $v = (2, 5, 4)$. Para encontrar proy (u, v), primero hallamos

$$u \cdot v = 2 - 10 + 12 = 4 \qquad \text{y} \qquad \|v\|^2 = 4 + 25 + 16 = 45$$

Entonces

$$\text{proy}\,(u, v) = \frac{u \cdot v}{\|v\|^2}\, v = \frac{4}{45}\,(2, 5, 4) = \left(\frac{8}{45}, \frac{20}{45}, \frac{16}{45}\right) = \left(\frac{8}{45}, \frac{4}{9}, \frac{16}{45}\right)$$

2.7. VECTORES LOCALIZADOS, HIPERPLANOS Y RECTAS EN $\mathbf{R}^n$

Esta sección distingue entre una n-pla $P(a_1, a_2, \ldots, a_n) \equiv P(a_i)$ vista como un punto de $\mathbf{R}^n$ y una n-pla $v = [c_1, c_2, \ldots, c_n]$ vista como un vector (flecha) desde el origen O hasta el punto $C[c_1, c_2, \ldots, c_n)$. Cualquier par de puntos $P = (a_i)$ y $Q = (b_i)$ en $\mathbf{R}^n$ define el *vector localizado* o *segmento dirigido* de P a Q, escrito $\overrightarrow{PQ}$. Identificamos $\overrightarrow{PQ}$ con el vector

$$v = Q - P = [b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n]$$

ya que $\overrightarrow{PQ}$ y v tienen la misma magnitud y dirección, tal y como se muestra en la Figura 2-6.

Un hiperplano H en $\mathbf{R}^n$ es el conjunto de puntos $(x_1, x_2, \ldots, x_n)$ que satisfacen una ecuación lineal no degenerada

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

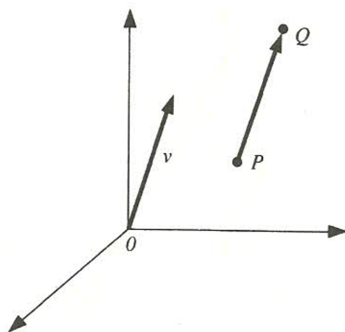


Figura 2-6.

En particular, un hiperplano H en $\mathbf{R}^2$ es una recta, y un hiperplano H en $\mathbf{R}^3$ es un plano. El vector $u = [a_1, a_2, \ldots, a_n] \neq 0$ recibe el nombre de *normal* a H. El término está justificado por el hecho (Problema 2.33) de que cualquier segmento dirigido $\overrightarrow{PQ}$, en el que P y Q pertenecen a H, es ortogonal al vector normal u. Esta propiedad, en el caso de $\mathbf{R}^3$, se muestra en la Figura 2-7.

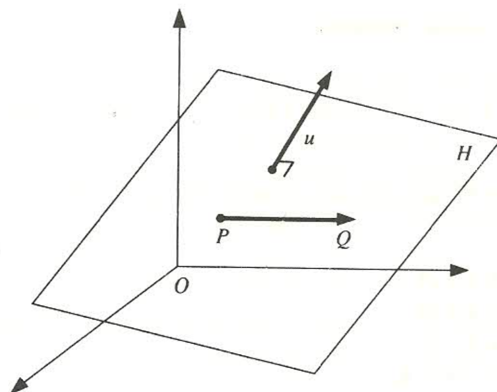


Figura 2-7.

La *recta* L en $\mathbf{R}^n$ que pasa por el punto $P(a_1, a_2, \ldots, a_n)$, en la dirección del vector no nulo $u = [u_1, u_2, \ldots, u_n]$, está constituida por los puntos $X(x_1, x_2, \ldots, x_n)$ que satisfacen

$$X = P + tu \quad \text{o} \quad \begin{cases} x_1 = a_1 + u_1 t \\ x_2 = a_2 + u_2 t \\ \ldots\ldots\ldots\ldots\ldots \\ x_n = a_n + u_n t \end{cases}$$

donde el *parámetro* t toma todos los valores reales. (Véase la Figura 2-8.)

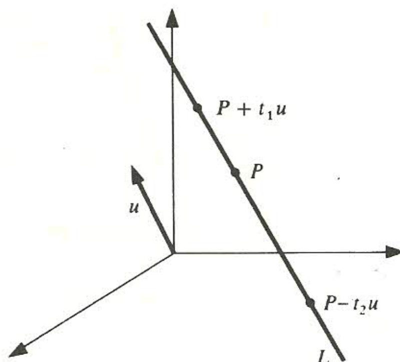


Figura 2-8.

EJEMPLO 2.8

a) Consideremos el hiperplano H en $\mathbf{R}^n$ que pasa por el punto $P(1, 3, -4, 2)$ y es normal al vector $u = [4, -2, 5, 6]$. Su ecuación tiene la forma

$$4x - 2y + 5z + 6t = k$$

Sustituyendo P en esta ecuación obtenemos

$$4(1) - 2(3) + 5(-4) + 6(2) = k \quad \text{o} \quad 4 - 6 - 20 + 12 = k \quad \text{o} \quad k = -10$$

Así $4x - 2y + 5z + 6t = -10$ es la ecuación de H.

b) Consideremos la recta L en $\mathbf{R}^4$ que pasa por el punto $P(1, 2, 3, -4)$, en la dirección de $u = [5, 6, -7, 8]$. Una representación paramétrica de L es la siguiente:

$$\begin{aligned} x_1 &= 1 + 5t \\ x_2 &= 2 + 6t \\ x_3 &= 3 - 7t \\ x_4 &= -4 + 8t \end{aligned} \qquad \text{o} \qquad (1 + 5t, 2 + 6t, 3 - 7t, -4 + 8t)$$

Nótese que $t = 0$ proporciona el punto P en L.

CURVAS EN $\mathbf{R}^n$

Sea D un intervalo (finito o infinito) en la recta real $\mathbf{R}$. Una función continua $F: D \to \mathbf{R}^n$ es una curva en $\mathbf{R}^n$. De este modo, a cada $t \in D$ se le asigna el siguiente punto (vector) en $\mathbf{R}^n$:

$$F(t) = [F_1(t), F_2(t), \ldots, F_n(t)]$$

Aún más, la derivada de $F(t)$ (si existe) proporciona el vector

$$V(t) = dF(t)/dt = [dF_1(t)/dt, dF_2(t)/dt, \ldots, dF_n(t)/dt]$$

que es tangente a la curva, y la normalización de $V(t)$ conduce a

$$T(t) = \frac{V(t)}{\|V(t)\|}$$

que es el vector unitario tangente a la curva. [Los vectores unitarios se denotan a menudo en negrita. Véase la Sección 2.8.]

EJEMPLO 2.9. Consideremos la siguiente curva C en $\mathbf{R}^3$:

$$F(t) = [\text{sen } t, \cos t, t]$$

Tomando la derivada de $F(t)$ [o de cada componente de $F(t)$] se obtiene

$$V(t) = [\cos t, -\text{sen } t, 1]$$

que es un vector tangente a la curva. Normalizamos $V(t)$. Primero obtenemos

$$\|V(t)\|^2 = \cos^2 t + \text{sen}^2 t + 1 = 1 + 1 = 2$$

Entonces

$$T(t) = \frac{V(t)}{\|V(t)\|} = \left[\frac{\cos t}{\sqrt{2}}, \frac{-\text{sen } t}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$$

que es el vector unitario tangente a la curva.

2.8. VECTORES ESPACIALES. NOTACION ijk EN $\mathbf{R}^3$

Los vectores en $\mathbf{R}^3$, denominados vectores espaciales, aparecen en numerosas aplicaciones, especialmente en física. De hecho, frecuentemente se utiliza una notación especial para tales vectores, que es la que se da a continuación:

$\mathbf{i} = (1, 0, 0)$ denota el vector unitario en la dirección x

$\mathbf{j} = (0, 1, 0)$ denota el vector unitario en la dirección y

$\mathbf{k} = (0, 0, 1)$ denota el vector unitario en la dirección z

En consecuencia, cualquier vector $u = (a, b, c)$ en $\mathbf{R}^3$ puede expresarse de forma única como sigue:

$$u = (a, b, c) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

Como $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ son vectores unitarios y son mutuamente ortogonales, tenemos

$$\mathbf{i}\cdot\mathbf{i} = 1, \quad \mathbf{j}\cdot\mathbf{j} = 1, \quad \mathbf{k}\cdot\mathbf{k} = 1 \quad \text{e} \quad \mathbf{i}\cdot\mathbf{j} = 0, \quad \mathbf{i}\cdot\mathbf{k} = 0, \quad \mathbf{j}\cdot\mathbf{k} = 0$$

Las diversas operaciones vectoriales discutidas previamente pueden expresarse en la presente notación como sigue. Supongamos $u = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ y $v = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$. Entonces

$$u + v = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k} \qquad u \cdot v = a_1 b_1 + a_2 b_2 + a_3 b_3$$

$$cu = ca_1\mathbf{i} + ca_2\mathbf{j} + ca_3\mathbf{k} \qquad \|u\| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

donde c es un escalar.

EJEMPLO 2.10. Considérense los vectores $u = 3\mathbf{i} + 5\mathbf{j} - 2\mathbf{k}$ y $v = 4\mathbf{i} - 3\mathbf{j} + 7\mathbf{k}$.

a) Para hallar $u + v$ sumamos las componentes correspondientes obteniendo

$$u + v = 7\mathbf{i} + 2\mathbf{j} + 5\mathbf{k}$$

b) Para hallar $3u - 2v$, primero multiplicamos los vectores por los escalares y después los sumamos:

$$3u - 2v = (9\mathbf{i} + 15\mathbf{j} - 6\mathbf{k}) + (-8\mathbf{i} + 6\mathbf{j} - 14\mathbf{k}) = 4\mathbf{i} + 21\mathbf{j} - 20\mathbf{k}$$

c) Para hallar $u \cdot v$ multiplicamos las componentes correspondientes sumándolas después:

$$u \cdot v = 12 - 15 - 14 = -17$$

d) Para hallar $\|u\|$ elevamos al cuadrado cada componente y después las sumamos obteniendo $\|u\|^2$. Esto es,

$$\|u\|^2 = 9 + 25 + 4 = 38 \qquad \text{y por tanto} \qquad \|u\| = \sqrt{38}$$

PRODUCTO VECTORIAL

Existe una operación especial para vectores u, v en $\mathbf{R}^3$, llamada *producto vectorial* y denotada por $u \times v$. Específicamente, supongamos

$$u = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \qquad \text{y} \qquad v = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

Entonces

$$u \times v = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$$

Nótese que $u \times v$ es un vector (de ahí su nombre). $u \times v$ también se denomina *producto externo* de u y v.

En notación de determinantes (Capítulo 7), con $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$, el producto vectorial puede expresarse también como sigue:

$$u \times v = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k}$$

o, equivalentemente,

$$u \times v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Dos importantes propiedades del producto vectorial son las siguientes (véase el Problema 2.56):

Teorema 2.5: Sean u, v y w vectores en $\mathbf{R}^3$.

i) El vector $w = u \times v$ es ortogonal a u y a v.

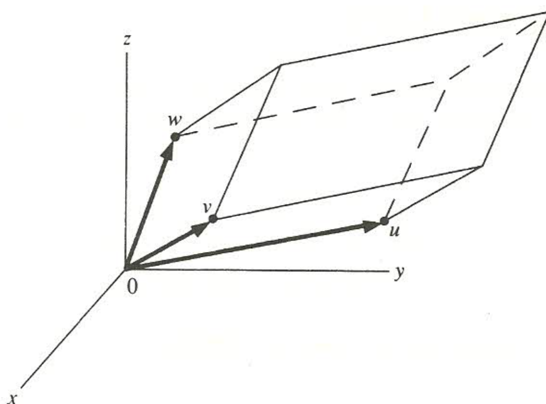


Figura 2-9.

ii) El valor absoluto del «producto triple» $u \cdot v \times w$ representa el volumen del paralelepípedo determinado por los tres vectores (como se muestra en la Figura 2-9).

EJEMPLO 2.11

a) Supongamos $u = 4\mathbf{i} + 3\mathbf{j} + 6\mathbf{k}$ y $v = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$. Entonces

$$u \times v = \begin{vmatrix} 3 & 6 \\ 5 & -3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 4 & 6 \\ 2 & -3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix} \mathbf{k} = -39\mathbf{i} + 24\mathbf{j} + 14\mathbf{k}$$

b)

$$(2, -1, 5) \times (3, 7, 6) = \left(\begin{vmatrix} -1 & 5 \\ 7 & 6 \end{vmatrix}, - \begin{vmatrix} 2 & 5 \\ 3 & 6 \end{vmatrix}, \begin{vmatrix} 2 & -1 \\ 3 & 7 \end{vmatrix} \right) = (-41, 3, 17)$$

(Aquí hallamos el producto vectorial sin utilizar la notación **ijk**.)

c) Los productos vectoriales de **i**, **j**, **k** son los siguientes:

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j} \quad \text{y} \quad \mathbf{j} \times \mathbf{i} = -\mathbf{k}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}$$

Dicho de otro modo, si vemos la terna (**i**, **j**, **k**) como una permutación cíclica, esto es, colocada sobre un círculo en el sentido contrario al de las agujas del reloj como en la Figura 2-10, entonces el producto de dos de los vectores en el sentido dado es el tercero, pero el producto de dos de ellos en el sentido contrario es el tercero con signo opuesto.

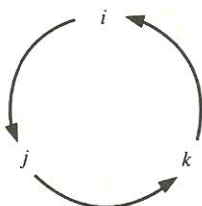


Figura 2-10.

2.9. NUMEROS COMPLEJOS

El conjunto de los números complejos se denota por **C**. Formalmente, un número complejo es un par ordenado (a, b) de números reales. La igualdad, suma y producto de números complejos se definen como sigue:

$$(a, b) = (c, d) \quad \text{si y sólo si} \quad a = c \quad \text{y} \quad b = d$$
$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b)(c, d) = (ac - bd, ad + bc)$$

Identificamos cada número real a con el número complejo $(a, 0)$:

$$a \leftrightarrow (a, 0)$$

Esto es factible debido a que las operaciones de suma y producto de números reales se conservan bajo dicha correspondencia

$$(a, 0) + (b, 0) = (a + b, 0) \quad \text{y} \quad (a, 0)(b, 0) = (ab, 0)$$

Así vemos **R** como un subconjunto de **C** y reemplazamos $(a, 0)$ por a cuando quiera que sea conveniente y posible.

El número complejo $(0, 1)$, denotado por i, tiene la importante propiedad de que

$$i^2 = ii = (0, 1)(0, 1) = (-1, 0) = -1 \quad \text{o} \quad i = \sqrt{-1}$$

Aún más, utilizando el hecho de que

$$(a, b) = (a, 0) + (0, b) \quad \text{y} \quad (0, b) = (b, 0)(0, 1)$$

tenemos

$$(a, b) = (a, 0) + (b, 0)(0, 1) = a + bi$$

La notación $z = a + bi$, donde $a \equiv \text{Re } z$ y $b \equiv \text{Im } z$ se denominan, respectivamente, las *partes real* e *imaginaria* del número complejo z, es más conveniente que (a, b). Por ejemplo, la suma y el producto de dos números complejos $z = a + bi$ y $w = c + di$ puede obtenerse simplemente utilizando las propiedades conmutativa y distributiva, además de $i^2 = -1$:

$$z + w = (a + bi) + (c + di) = a + c + bi + di = (a + c) + (b + d)i$$
$$zw = (a + bi)(c + di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i$$

Advertencia: El uso anterior de la letra i para denotar $\sqrt{-1}$ no tiene, en ningún caso, relación con la notación vectorial $\mathbf{i} = (1, 0, 0)$ introducida en la Sección 2.8.

El conjugado del número complejo $z = (a, b) = a + bi$ se denota y define por

$$\bar{z} = \overline{a + bi} = a - bi$$

Entonces $z\bar{z} = (a + bi)(a - bi) = a^2 - b^2 i^2 = a^2 + b^2$. Si, además, $z \neq 0$, entonces el inverso z^{-1} de z y la división de w por z vienen dados, respectivamente, por

$$z^{-1} = \frac{\bar{z}}{z\bar{z}} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i \qquad \text{y} \qquad \frac{w}{z} = wz^{-1}$$

donde $w \in \mathbf{C}$. También definimos

$$-z = -1z \qquad \text{y} \qquad w - z = w + (-z)$$

Así como los números reales pueden representarse por los puntos de una recta, los números complejos pueden representarse por los puntos del plano. Concretamente, el punto (a, b) del plano representará el número complejo $z = a + bi$, es decir, el número cuya parte real es a y cuya parte imaginaria es b. (Véase la Figura 2-11.) El *valor absoluto* de z, escrito $|z|$, se define como la distancia de z al origen:

$$|z| = \sqrt{a^2 + b^2}$$

Nótese que $|z|$ es igual a la norma del vector (a, b). Además, $|z| = \sqrt{z\bar{z}}$.

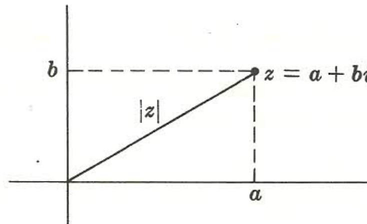


Figura 2-11.

EJEMPLO 2.12. Considérense los vectores $z = 2 + 3i$ y $w = 5 - 2i$. Entonces

$$z + w = (2 + 3i) + (5 - 2i) = 2 + 5 + 3i - 2i = 7 + i$$

$$zw = (2 + 3i)(5 - 2i) = 10 + 15i - 4i - 6i^2 = 16 + 11i$$

$$\bar{z} = \overline{2 + 3i} = 2 - 3i \qquad \text{y} \qquad \bar{w} = \overline{5 - 2i} = 5 + 2i$$

$$\frac{w}{z} = \frac{5 - 2i}{2 + 3i} = \frac{(5 - 2i)(2 - 3i)}{(2 + 3i)(2 - 3i)} = \frac{4 - 19i}{13} = \frac{4}{13} - \frac{19}{13} i$$

$$|z| = \sqrt{4 + 9} = \sqrt{13} \qquad \text{y} \qquad |w| = \sqrt{25 + 4} = \sqrt{29}$$

Nota: En el Apéndice definimos la estructura algebraica denominada *cuerpo*. Es oportuno enfatizar que el conjunto **C** de los números complejos, con las operaciones de suma y producto anteriores, es un cuerpo.

2.10. VECTORES EN $\mathbf{C}^n$

El conjunto de todas las n-plas de números complejos, denotado por $\mathbf{C}^n$, se llama *n-espacio complejo*. Tal como en el caso real, los elementos de $\mathbf{C}^n$ se llaman *puntos* o *vectores* y los de $\mathbf{C}$, *escalares*. La *suma vectorial* y el *producto por un escalar* en $\mathbf{C}^n$ vienen dados por

$$(z_1, z_2, \ldots, z_n) + (w_1, w_2, \ldots, w_n) = (z_1 + w_1, z_2 + w_2, \ldots, z_n + w_n)$$
$$z(z_1, z_2, \ldots, z_n) = (zz_1, zz_2, \ldots, zz_n)$$

donde z_i, w_i, $z \in \mathbf{C}$.

EJEMPLO 2.13

a) $(2 + 3i, 4 - i, 3) + (3 - 2i, 5i, 4 - 6i) = (5 + i, 4 + 4i, 7 - 6i)$.

b) $2i(2 + 3i, 4 - i, 3) = (-6 + 4i, 2 + 8i, 6i)$.

Ahora sean u y v vectores arbitrarios en $\mathbf{C}^n$:

$$u = (z_1, z_2, \ldots, z_n) \qquad v = (w_1, w_2, \ldots, w_n) \qquad z_i, w_i \in \mathbf{C}$$

El *producto escalar* o *interno* de u y v se define como sigue:

$$u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

Nótese que esta definición se reduce a la dada previamente en el caso real, puesto que $w_i = \bar{w}_i$ cuando w_i es real. La norma de u se define como

$$\|u\| = \sqrt{u \cdot u} = \sqrt{z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

Obsérvese que $u \cdot u$ y por tanto $\|u\|$ son reales y positivos cuando $u \neq 0$ y 0 cuando $u = 0$.

EJEMPLO 2.14. Sean $u = (2 + 3i, 4 - i, 2i)$ y $v = (3 - 2i, 5, 4 - 6i)$. Entonces

$$\begin{aligned} u \cdot v &= (2 + 3i)\overline{(3 - 2i)} + (4 - i)\overline{(5)} + (2i)\overline{(4 - 6i)} = \\ &= (2 + 3i)(3 + 2i) + (4 - i)(5) + (2i)(4 + 6i) = \\ &= 13i + 20 - 5i - 12 + 8i = 8 + 16i \end{aligned}$$

$$\begin{aligned} u \cdot u &= (2 + 3i)\overline{(2 + 3i)} + (4 - i)\overline{(4 - i)} + (2i)\overline{(2i)} = \\ &= (2 + 3i)(2 - 3i) + (4 - i)(4 + i) + (2i)(-2i) = \\ &= 13 + 17 + 4 = 34 \end{aligned}$$

$$\|u\| = \sqrt{u \cdot u} = \sqrt{34}$$

El espacio $\mathbf{C}^n$, con las operaciones anteriores de suma vectorial, producto por un escalar y producto escalar, se denomina *n-espacio euclídeo complejo*.

Nota: Si $u \cdot v$ estuviera definido por $u \cdot v = z_1 w_1 + \cdots + z_n w_n$, podría darse $u \cdot u = 0$ incluso a pesar de ser $u \neq 0$, como sucedería, por ejemplo, si fuera $u = (1, i, 0)$. De hecho, $u \cdot u$ podría no ser siquiera real.

PROBLEMAS RESUELTOS

VECTORES EN $\mathbf{R}^n$

2.1. Sean $u = (2, -7, 1)$, $v = (-3, 0, 4)$, $w = (0, 5, -8)$. Hallar *a)* $3u - 4v$, *b)* $2u + 3v - 5w$.

Primero efectuamos el producto por los escalares y luego la suma de vectores.

a) $3u - 4v = 3(2, -7, 1) - 4(-3, 0, 4) = (6, -21, 3) + (12, 0, -16) = (18, -21, -13)$.

b) $2u + 3v - 5w = 2(2, -7, 1) + 3(-3, 0, 4) - 5(0, 5, -8) =$
$= (4, -14, 2) + (-9, 0, 12) + (0, -25, 40) =$
$= (4 - 9 + 0, -14 + 0 - 25, 2 + 12 + 40) = (-5, -39, 54)$.

2.2. Calcular:

$$a)\quad 2\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} - 3\begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}, \qquad b)\quad -2\begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix} + 4\begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix} - 3\begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix}$$

Primero efectuamos el producto por los escalares y luego la suma de vectores.

$$a)\quad 2\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} - 3\begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 6 \end{pmatrix} + \begin{pmatrix} -6 \\ -9 \\ 12 \end{pmatrix} = \begin{pmatrix} -4 \\ -11 \\ 18 \end{pmatrix}$$

$$b)\quad -2\begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix} + 4\begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix} - 3\begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -10 \\ -6 \\ 8 \end{pmatrix} + \begin{pmatrix} -4 \\ 20 \\ 8 \end{pmatrix} + \begin{pmatrix} -9 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -23 \\ 17 \\ 19 \end{pmatrix}$$

2.3. Hallar x e y si *a)* $(x, 3) = (2, x + y)$; *b)* $(4, y) = x(2, 3)$.

a) Siendo iguales los dos vectores, las componentes correspondientes son iguales entre sí:

$$x = 2 \qquad 3 = x + y$$

Sustituimos $x = 2$ en la segunda ecuación obteniendo $y = 1$. Así $x = 2$ e $y = 1$.

b) Multiplicamos por el escalar x para obtener $(4, y) = x(2, 3) = (2x, 3x)$. Igualamos entre sí las componentes correspondientes:

$$4 = 2x \qquad y = 3x$$

Resolvemos ahora las ecuaciones lineales para x e y: $x = 2$ e $y = 6$.

2.4. Probar el Teorema 2.1.

Sean u_i, v_i y w_i las componentes i-ésimas de u, v y w, respectivamente.

i) Por definición, $u_i + v_i$ es la componente i-ésima de $u + v$ y por tanto $(u_i + v_i) + w_i$ es la de $(u + v) + w$. Por otra parte, $v_i + w_i$ es la componente i-ésima de $v + w$ y en consecuencia

$u_i + (v_i + w_i)$ es la de $u + (v + w)$. Pero u_i, v_i y w_i son números reales para los que se verifica la ley asociativa, esto es,

$$(u_i + v_i) + w_i = u_i + (v_i + w_i) \qquad \text{para} \qquad i = 1, \ldots, n$$

De acuerdo con esto, $(u + v) + w = u + (v + w)$ ya que sus componentes correspondientes son iguales.

ii) Aquí $0 = (0, 0, \ldots, 0)$; por tanto,

$$u + 0 = (u_1, u_2, \ldots, u_n) + (0, 0, \ldots, 0) =$$
$$= (u_1 + 0, u_2 + 0, \ldots, u_n + 0) = (u_1, u_2, \ldots, u_n) = u$$

iii) Dado que $-u = -1(u_1, u_2, \ldots, u_n) = (-u_1, -u_2, \ldots, -u_n)$,

$$u + (-u) = (u_1, u_2, \ldots, u_n) + (-u_1, -u_2, \ldots, -u_n) =$$
$$= (u_1 - u_1, u_2 - u_2, \ldots, u_n - u_n) = (0, 0, \ldots, 0) = 0$$

iv) Por definición, $u_i + v_i$ es la componente i-ésima de $u + v$ y $v_i + u_i$ es la de $v + u$. Pero u_i y v_i son números reales para los que se verifica la ley conmutativa, esto es,

$$u_i + v_i = v_i + u_i \qquad i = 1, \ldots, n$$

Por tanto, $u + v = v + u$, puesto que sus componentes correspondientes son iguales.

v) Dado que $u_i + v_i$ es la componente i-ésima de $u + v$, $k(u_i + v_i)$ es la de $k(u + v)$. Como ku_i y kv_i son las componentes i-ésimas de ku y kv, respectivamente, $ku_i + kv_i$ es la de $ku + kv$. Pero k, u_i y v_i son números reales. Por tanto,

$$k(u_i + v_i) = ku_i + kv_i \qquad i = 1, \ldots, n$$

Así $k(u + v) = ku + kv$, porque sus componentes correspondientes son iguales.

vi) Obsérvese que el primer signo más se refiere a la suma de dos escalares k y k', mientras que el segundo se refiere a la suma de dos vectores ku y $k'u$.

Por definición, $(k + k')u_i$ es la componente i-ésima del vector $(k + k')u$. Dado que ku_i y $k'u_i$ son las componentes i-ésimas de ku y $k'u$, respectivamente, $ku_i + k'u_i$ es la del vector $ku + k'u$. Pero k, k' y u_i son números reales. Entonces

$$(k + k')u_i = ku_i + k'u_i \qquad i = 1, \ldots, n$$

De este modo, $(k + k')u = ku + k'u$ ya que sus componentes correspondientes son iguales.

vii) Como $k'u_i$ es la componente i-ésima de $k'u$, $k(k'u_i)$ es la de $k(k'u)$. Pero $(kk')u_i$ es la componente i-ésima de $(kk')u$ y, debido a que k, k' y u_i son números reales,

$$(kk')u_i = k(k'u_i) \qquad i = 1, \ldots, n$$

Por consiguiente, $(kk')u = k(k'u)$, puesto que sus componentes correspondientes son iguales.

viii) $1 \cdot u = 1(u_1, u_2, \ldots, u_n) = (1u_1, 1u_2, \ldots, 1u_n) = (u_1, u_2, \ldots, u_n) = u$.

VECTORES Y ECUACIONES LINEALES

2.5. Convertir la siguiente ecuación vectorial en un sistema de ecuaciones lineales equivalente y resolverlo:

$$\begin{pmatrix} 1 \\ -6 \\ 5 \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} + z\begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$$

Multiplicamos los vectores de la derecha por los escalares desconocidos y luego los sumamos:

$$\begin{pmatrix} 1 \\ -6 \\ 5 \end{pmatrix} = \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix} + \begin{pmatrix} 2y \\ 5y \\ 8y \end{pmatrix} + \begin{pmatrix} 3z \\ 2z \\ 3z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 2x + 5y + 2z \\ 3x + 8y + 3z \end{pmatrix}$$

Igualamos entre sí las componentes correspondientes de los vectores y reducimos el sistema a forma escalonada:

$$\begin{aligned} x + 2y + 3z &= 1 \\ 2x + 5y + 2z &= -6 \\ 3x + 8y + 3z &= 5 \end{aligned} \qquad \begin{aligned} x + 2y + 3z &= 1 \\ y - 4z &= -8 \\ 2y - 6z &= 2 \end{aligned} \qquad \begin{aligned} x + 2y + 3z &= 1 \\ y - 4z &= -8 \\ 2z &= 18 \end{aligned}$$

El sistema es triangular, y por sustitución hacia atrás se obtiene la solución única: $x = -82$, $y = 28$, $z = 9$.

2.6. Escribir el vector $v = (1, -2, 5)$ como combinación lineal de los vectores $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 3)$ y $u_3 = (2, -1, 1)$.

Queremos expresar v en la forma $v = xu_1 + yu_2 + zu_3$, con x, y y z aún por determinar. Siendo así tenemos

$$\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + z\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} x + y + 2z \\ x + 2y - z \\ x + 3y + z \end{pmatrix}$$

(Para formar combinaciones lineales, resulta más conveniente escribir los vectores como columnas que como filas.) Igualando las componentes correspondientes entre sí obtenemos:

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned} \quad \text{o} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \quad \text{o} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

La solución única del sistema en forma triangular es $x = -6$, $y = 3$, $z = 2$; así $v = -6u_1 + 3u_2 + 2u_3$.

2.7. Escribir el vector $v = (2, 3, -5)$ como combinación lineal de $u_1 = (1, 2, -3)$, $u_2 = (2, -1, -4)$ y $u_3 = (1, 7, -5)$.

Hallamos el sistema de ecuaciones equivalentes y luego lo resolvemos. En primer lugar tomamos:

$$\begin{pmatrix} 2 \\ 3 \\ -5 \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + y\begin{pmatrix} 2 \\ -1 \\ -4 \end{pmatrix} + z\begin{pmatrix} 1 \\ 7 \\ -5 \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ 2x - y + 7z \\ -3x - 4y - 5z \end{pmatrix}$$

Igualando entre sí las componentes correspondientes obtenemos

$$\begin{aligned} x + 2y + z &= 2 \\ 2x - y + 7z &= 3 \\ -3x - 4y - 5z &= -5 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + z &= 2 \\ -5y + 5z &= -1 \\ 2y - 2z &= 1 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + z &= 2 \\ -5y + 5z &= -1 \\ 0 &= 3 \end{aligned}$$

La tercera ecuación, $0 = 3$, indica que el sistema no tiene solución. Entonces v no puede escribirse como combinación lineal de los vectores u_1, u_2 y u_3.

2.8. Determinar si los vectores $u_1 = (1, 1, 1)$, $u_2 = (2, -1, 3)$ y $u_3 = (1, -5, 3)$ son linealmente dependientes o linealmente independientes.

Recuérdese que u_1, u_2, u_3 son linealmente dependientes o linealmente independientes según que la ecuación vectorial $xu_1 + yu_2 + zu_3$ tenga o no una solución no nula. Así pues, para empezar igualamos una combinación lineal de los vectores al vector nulo:

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}2\\-1\\3\end{pmatrix} + z\begin{pmatrix}1\\-5\\3\end{pmatrix} = \begin{pmatrix}x + 2y + z\\x - y - 5z\\x + 3y + 3z\end{pmatrix}$$

Igualamos las componentes correspondientes entre sí y reducimos el sistema resultante a forma escalonada:

$$\begin{aligned} x + 2y + z &= 0\\ x - y - 5z &= 0\\ x + 3y + 3z &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + z &= 0\\ -3y - 6z &= 0\\ y + 2z &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + z &= 0\\ y + 2z &= 0 \end{aligned}$$

El sistema en forma escalonada tiene una variable libre; por consiguiente, el sistema tiene una solución no trivial. Siendo así, los vectores originales son linealmente dependientes. (No necesitamos resolver el sistema para determinar la dependencia o independencia lineal; basta conocer si existe una solución no nula.)

2.9. Determinar si los vectores $(1, -2, -3)$, $(2, 3, -1)$ y $(3, 2, 1)$ son linealmente dependientes.

Igualamos al vector cero una combinación lineal (con coeficientes x, y, z) de los vectores:

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = x\begin{pmatrix}1\\-2\\-3\end{pmatrix} + y\begin{pmatrix}2\\3\\-1\end{pmatrix} + z\begin{pmatrix}3\\2\\1\end{pmatrix} = \begin{pmatrix}x + 2y + 3z\\-2x + 3y + 2z\\-3x - y + z\end{pmatrix}$$

Igualamos las componentes correspondientes entre sí y reducimos el sistema resultante a forma escalonada:

$$\begin{aligned} x + 2y + 3z &= 0\\ -2x + 3y + 2z &= 0\\ -3x - y + z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + 3z &= 0\\ 7y + 8z &= 0\\ 5y + 10z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + 3z &= 0\\ y + 2z &= 0\\ 7y + 8z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + 3z &= 0\\ y + 2z &= 0\\ -6z &= 0 \end{aligned}$$

El sistema homogéneo está en forma triangular, sin variables libres; por tanto, sólo tiene la solución trivial. Así que los vectores originales son linealmente independientes.

2.10. Demostrar la siguiente afirmación: $n + 1$ o más vectores arbitrarios en $\mathbf{R}^n$ son linealmente dependientes.

Supongamos que $u_1, u_2, \ldots, u_q$ son vectores en $\mathbf{R}^n$ con $q > n$. La ecuación vectorial

$$x_1u_1 + x_2u_2 + \cdots + x_qu_q = 0$$

es equivalente a un sistema homogéneo de n ecuaciones con $q > n$ incógnitas. De acuerdo con el Teorema 1.9, este sistema tiene solución no trivial. Consecuentemente, $u_1, u_2, \ldots, u_q$ son linealmente dependientes.

2.11. Demostrar que cualquier conjunto de q vectores que contenga el vector cero es linealmente dependiente.

Denotando los vectores por $0, u_2, u_3, \ldots, u_q$ tenemos $1(0) + 0u_2 + 0u_3 + \cdots + 0u_q = 0$.

PRODUCTO INTERNO. ORTOGONALIDAD

2.12. Calcular $u \cdot v$, donde $u = (1, -2, 3, -4)$ y $v = (6, 7, 1, -2)$.

Multiplicamos las componentes correspondientes y sumamos:

$$u \cdot v = (1)(6) + (-2)(7) + (3)(1) + (-4)(-2) = 3$$

2.13. Sean $u = (3, 2, 1)$, $v = (5, -3, 4)$, $w = (1, 6, -7)$. Hallar: *a*) $(u + v) \cdot w$, *b*) $u \cdot w + v \cdot w$.

a) Primero calculamos $u + v$ sumando las componentes correspondientes:

$$u + v = (3 + 5, 2 - 3, 1 + 4) = (8, -1, 5)$$

Luego calculamos

$$(u + v) \cdot w = (8)(1) + (-1)(6) + (5)(-7) = 8 - 6 - 35 = -33$$

b) Primero hallamos $u \cdot w = 3 + 12 - 7 = 8$ y $v \cdot w = 5 - 18 - 28 = -41$. Entonces $u \cdot w + v \cdot w = 8 - 41 = -33$. [Tal y como cabía esperar, por el Teorema 2.2 i), ambos valores son iguales.]

2.14. Sean $u = (1, 2, 3, -4)$, $v = (5, -6, 7, 8)$ y $k = 3$. Hallar: *a*) $k(u \cdot v)$, *b*) $(ku) \cdot v$, *c*) $u \cdot (kv)$.

a) Primero determinamos $u \cdot v = 5 - 12 + 21 - 32 = -18$. Entonces $k(u \cdot v) = 3(-18) = -54$.

b) Primero determinamos $ku = (3(1), 3(2), 3(3), 3(-4)) = (3, 6, 9, -12)$. Entonces

$$(ku) \cdot v = (3)(5) + (6)(-6) + (9)(7) + (-12)(8) = 15 - 36 + 63 - 96 = -54$$

c) Primero determinamos $kv = (15, -18, 21, 24)$. Entonces

$$u \cdot (kv) = (1)(15) + (2)(-18) + (3)(21) + (-4)(24) = 15 - 36 + 63 - 96 = -54$$

2.15. Sean $u = (5, 4, 1)$, $v = (3, -4, 1)$ y $w = (1, -2, 3)$. ¿Qué pares de dichos vectores son perpendiculares?

Hallamos el producto escalar de cada par de vectores:

$$u \cdot v = 15 - 16 + 1 = 0 \qquad v \cdot w = 3 + 8 + 3 = 14 \qquad u \cdot w = 5 - 8 + 3 = 0$$

Por consiguiente, tanto u y v como u y w son ortogonales, pero v y w no lo son.

2.16. Determinar el valor de k para que los vectores u y v sean ortogonales, siendo $u = (1, k, -3)$ y $v = (2, -5, 4)$.

Calculamos $u \cdot v$, lo igualamos a 0 y resolvemos para k. $u \cdot v = (1)(2) + (k)(-5) + (-3)(4) = 2 - 5k - 12 = 0$ o $-5k - 10 = 0$. Resolviendo, $k = -2$.

2.17. Demostrar el Teorema 2.2.

Sean $u = (u_1, u_2, \ldots, u_n)$, $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n)$.

i) Dado que $u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$,

$$\begin{aligned}(u + v) \cdot w &= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \cdots + (u_n + v_n)w_n = \\ &= u_1w_1 + v_1w_1 + u_2w_2 + v_2w_2 + \cdots + u_nw_n + v_nw_n = \\ &= (u_1w_1 + u_2w_2 + \cdots + u_nw_n) + (v_1w_1 + v_2w_2 + \cdots + v_nw_n) = \\ &= u \cdot w + v \cdot w\end{aligned}$$

ii) Como $ku = (ku_1, ku_2, \ldots, ku_n)$,

$$(ku) \cdot v = ku_1 v_1 + ku_2 v_2 + \cdots + ku_n v_n = k(u_1 v_1 + u_2 v_2 + \cdots + u_n v_n) = k(u \cdot v)$$

iii) $u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n = v \cdot v.$

iv) Como u_i^2 es no negativo para cada i, y dado que la suma de números reales no negativos es no negativa,

$$u \cdot u = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$$

Además, $u \cdot u = 0$ si y sólo si $u_i = 0$ para todo i, esto es, si y sólo si $u = 0$.

NORMA (LONGITUD) EN $\mathbf{R}^n$

2.18. Hallar $\|w\|$ si $w = (-3, 1, -2, 4, -5)$.

$\|w\|^2 = (-3)^2 + 1^2 + (-2)^2 + 4^2 + (-5)^2 = 9 + 1 + 4 + 16 + 25 = 55$; de aquí $\|w\| = \sqrt{55}$.

2.19. Determinar el valor de k para que $\|u\| = \sqrt{39}$, donde $u = (1, k, -2, 5)$.

$\|u\|^2 = 1^2 + k^2 + (-2)^2 + 5^2 = k^2 + 30$. Ahora resolvemos $k^2 + 30 = 39$ y obtenemos $k = 3, -3$.

2.20. Normalizar $w = (4, -2, -3, 8)$.

Primero hallamos $\|w\|^2 = w \cdot w = 4^2 + (-2)^2 + (-3)^2 + 8^2 = 16 + 4 + 9 + 64 = 93$. Dividimos cada componente de w por $\|w\| = \sqrt{93}$ obteniendo

$$\hat{w} = \frac{w}{\|w\|} = \left(\frac{4}{\sqrt{93}}, \frac{-2}{\sqrt{93}}, \frac{-3}{\sqrt{93}}, \frac{8}{\sqrt{93}}\right)$$

2.21. Normalizar $v = \left(\frac{1}{2}, \frac{2}{3}, -\frac{1}{4}\right)$.

Nótese que tanto v como cualquiera de sus múltiplos positivos tendrán la misma forma normalizada. En tal caso, podemos multiplicar primero v por 12 para eliminar los denominadores: $12v = (6, 8, -3)$. Entonces

$$\|12v\|^2 = 36 + 64 + 9 = 109 \quad \text{y} \quad \hat{v} = \widehat{12v} = \frac{12v}{\|12v\|} = \left(\frac{6}{\sqrt{109}}, \frac{8}{\sqrt{109}}, \frac{-3}{\sqrt{109}}\right)$$

2.22. Demostrar el Teorema 2.3 (Cauchy-Schwarz).

En su lugar probaremos el siguiente resultado, más potente: $|u \cdot v| \leq \sum_{i=1}^{n} |u_i v_i| \leq \|u\| \|v\|$. Para empezar, si $u = 0$ o $v = 0$, la desigualdad se reduce a $0 \leq 0 \leq 0$ y es, por tanto, cierta. En consecuencia, sólo necesitamos considerar el caso en el que $u \neq 0$ y $v \neq 0$, esto es, en el que $\|u\| \neq 0$ y $\|v\| \neq 0$. Además, debido a que

$$|u \cdot v| = |\sum u_i v_i| \leq \sum |u_i v_i|$$

sólo es necesario probar la segunda desigualdad.

Ahora bien, para dos números reales cualesquiea $x, y \in \mathbf{R}$, $0 \leq (x - y)^2 = x^2 - 2xy + y^2$ o, equivalentemente,

$$2xy \leq x^2 + y^2 \qquad [1]$$

Tomamos $x = |u_i|/\|u\|$ e $y = |v_i|/\|v\|$ en [1] para obtener, para cualquier i,

$$2\frac{|u_i|}{\|u\|}\frac{|v_i|}{\|v\|} \le \frac{|u_i|^2}{\|u\|^2} + \frac{|v_i|^2}{\|v\|^2} \qquad [2]$$

Pero, por definición de norma de un vector, $\|u\| = \sum u_i^2 = \sum |u_i|^2$ y $\|v\| = \sum v_i^2 = \sum |v_i|^2$. Así, sumando en [2] respecto a i y utilizando $|u_i v_i| = |u_i|\,|v_i|$, tenemos

$$2\frac{\sum |u_i v_i|}{\|u\|\,\|v\|} \le \frac{\sum |u_i|^2}{\|u\|^2} + \frac{\sum |v_i|^2}{\|v\|^2} = \frac{\|u\|^2}{\|u\|^2} + \frac{\|v\|^2}{\|v\|^2} = 2$$

esto es,

$$\frac{\sum |u_i v_i|}{\|u\|\,\|v\|} \le 1$$

Multiplicando a ambos lados por $\|u\|\,\|v\|$ obtenemos la desigualdad pedida.

2.23. Probar el Teorema 2.4 (Minkowski).

Por la desigualdad de Cauchy-Schwarz (Problema 2.22) y el resto de las propiedades del producto interno,

$$\|u + v\|^2 = (u + v)\cdot(u + v) = u\cdot u + 2(u\cdot v) + v\cdot v \le$$
$$\le \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

Tomando las raíces cuadradas de ambos miembros se obtiene la desigualdad buscada.

2.24. Probar que la norma en $\mathbf{R}^n$ satisface las siguientes leyes:

a) $[N_1]$ Para cualquier vector u, $\|u\| \ge 0$; y $\|u\| = 0$ si y sólo si $u = 0$.

b) $[N_2]$ Para cualquier vector u y cualquier escalar k, $\|ku\| = |k|\,\|u\|$.

c) $[N_3]$ Para dos vectores cualesquiera u y v, $\|u + v\| \le \|u\| + \|v\|$.

a) Según el Teorema 2.2, $u\cdot u \ge 0$ y $u\cdot u = 0$ si y sólo si $u = 0$. Como $\|u\| = \sqrt{u\cdot u}$, resulta $[N_1]$.

b) Supongamos $u = (u_1, u_2, \ldots, u_n)$ y por tanto $ku = (ku_1, ku_2, \ldots, ku_n)$. Entonces

$$\|ku\|^2 = (ku_1)^2 + (ku_2)^2 + \cdots + (ku_n)^2 = k^2(u_1^2 + u_2^2 + \cdots + u_n^2) = k^2\|u\|^2$$

Tomando raíces cuadradas se obtiene $[N_2]$.

c) $[N_3]$ se ha demostrado en el Problema 2.23.

2.25. Sean $u = (1, 2, -2)$, $v = (3, -12, 4)$ y $k = -3$. *a*) Encontrar $\|u\|$, $\|v\|$ y $\|ku\|$. *b*) Comprobar que $\|ku\| = |k|\,\|u\|$ y $\|u + v\| \le \|u\| + \|v\|$.

a) $\|u\| = \sqrt{1 + 4 + 4} = \sqrt{9} = 3$, $\|v\| = \sqrt{9 + 144 + 16} = \sqrt{169} = 13$, $ku = (-3, -6, 6)$ y $\|ku\| = \sqrt{9 + 36 + 36} = \sqrt{81} = 9$.

b) Dado que $|k| = |-3| = 3$, tenemos $|k|\,\|u\| = 3\cdot 3 = 9 = \|ku\|$. Además $u + v = (4, -10, 2)$. Así

$$\|u + v\| = \sqrt{16 + 100 + 4} = \sqrt{120} \le 16 = 3 + 13 = \|u\| + \|v\|$$

DISTANCIA. ANGULOS. PROYECCIONES

2.26. Hallar la distancia $d(u, v)$ entre los vectores u y v, donde

a) $u = (1, 7), v = (6, -5)$.

b) $u = (3, -5, 4), v = (6, 2, -1)$.

c) $u = (5, 3, -2, -4, -1), v = (2, -1, 0, -7, 2)$.

En cada caso utilizamos la fórmula $d(u, v) = \|u - v\| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}$.

a) $d(u, v) = \sqrt{(1-6)^2 + (7+5)^2} = \sqrt{25 + 144} = \sqrt{169} = 13$.

b) $d(u, v) = \sqrt{(3-6)^2 + (-5-2)^2 + (4+1)^2} = \sqrt{9 + 49 + 25} = \sqrt{13}$.

c) $d(u, v) = \sqrt{(5-2)^2 + (3+1)^2 + (-2+0)^2 + (-4+7)^2 + (-1-2)^2} = \sqrt{47}$.

2.27. Encontrar un número k tal que $d(u, v) = 6$, siendo $u = (2, k, 1, -4)$ y $v = (3, -1, 6, -3)$.

Primero hallamos

$$[d(u, v)]^2 = (2-3)^2 + (k+1)^2 + (1-6)^2 + (-4+3)^2 = k^2 + 2k + 28$$

Ahora resolvemos $k^2 + 2k + 28 = 6^2$ obteniendo $k = 2, -4$.

2.28. A partir del Problema 2.24, demostrar que la función distancia satisface:

$[M_1]$ $d(u, v) \geq 0$; y $d(u, v) = 0$ si y sólo si $u = v$.

$[M_2]$ $d(u, v) = d(v, u)$.

$[M_3]$ $d(u, v) \leq d(u, w) + d(w, v)$ (desigualdad triangular).

$[M_1]$ es consecuencia directa de $[N_1]$. Según $[N_2]$,

$$d(u, v) = \|u - v\| = \|(-1)(v - u)\| = |-1|\,\|v - u\| = \|v - u\| = d(v, u)$$

que es $[M_2]$. Según $[N_3]$,

$$d(u, v) = \|u - v\| = \|(u - w) + (w - v)\| \leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v)$$

que es $[M_3]$.

2.29. Determinar $\cos\theta$, donde θ es el ángulo entre $u = (1, 2, -5)$ y $v = (2, 4, 3)$.

Primero hallamos

$$u \cdot v = 2 + 8 - 15 = -5 \qquad \|u\|^2 = 1 + 4 + 25 = 30 \qquad \|v\|^2 = 4 + 16 + 9 = 29$$

Luego

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|} = -\frac{5}{\sqrt{30}\sqrt{29}}$$

2.30. Determinar proy (u, v), donde $u = (1, -3, 4)$ y $v = (3, 4, 7)$.

Primero hallamos $u \cdot v = 3 - 12 + 28 = 19$ y $\|v\|^2 = 9 + 16 + 49 = 74$. Luego

$$\text{proy}(u, v) = \frac{u \cdot v}{\|v\|^2} v = \frac{19}{74}(3, 4, 7) = \left(\frac{57}{74}, \frac{76}{74}, \frac{133}{74}\right) = \left(\frac{57}{74}, \frac{38}{37}, \frac{133}{74}\right)$$

PUNTOS, RECTAS E HIPERPLANOS

En esta sección se distingue entre una n-pla $P(a_1, a_2, \ldots, a_n) \equiv P(a_i)$ vista como un punto en $\mathbf{R}^n$ y una n-pla $v = [c_1, c_2, \ldots, c_n]$ vista como un vector (flecha) desde el origen O hasta el punto $C(c_1, c_2, \ldots, c_n)$.

2.31. Hallar el vector v que se identifica con el segmento dirigido $\overrightarrow{PQ}$ para los puntos *a*) $P(2, 5)$ y $Q(-3, 4)$ en $\mathbf{R}^2$, *b*) $P(1, -2, 4)$ y $Q(6, 0, -3)$ en $\mathbf{R}^3$.

a) $v = Q - P = [-3 - 2, 4 - 5] = [-5, -1]$.

b) $v = Q - P = [6 - 1, 0 + 2, -3 - 4] = [5, 2, -7]$.

2.32. Considérense los puntos $P(3, k, -2)$ y $Q(5, 3, 4)$ en $\mathbf{R}^3$. Encontrar un valor de k para el que $\overrightarrow{PQ}$ sea ortogonal al vector $u = [4, -3, 2]$.

Primero hallamos $v = Q - P = [5 - 3, 3 - k, 4 + 2] = [2, 3 - k, 6]$. A continuación calculamos

$$u \cdot v = 4 : 2 - 3(3 - k) + 2 \cdot 6 = 8 - 9 + 3k + 12 = 3k + 11$$

Por último, imponemos $u \cdot v = 0$ ó $3k + 11 = 0$, de donde $k = -11/3$.

2.33. Considérese el hiperplano H en $\mathbf{R}^n$ identificable con el conjunto solución de la ecuación lineal

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b \qquad [1]$$

donde $u = [a_1, a_2, \ldots, a_n] \neq 0$. Demostrar que el segmento dirigido $\overrightarrow{PQ}$ para cualquier par de puntos $P, Q \in H$ es ortogonal al vector de coeficientes u. Se dice que el vector u es *normal* al hiperplano H.

Sean $w_1 = \overrightarrow{OP}$ y $w_2 = \overrightarrow{OQ}$; por consiguiente, $v = w_2 - w_1 = \overrightarrow{PQ}$. Por [1], $u \cdot w_1 = b$ y $u \cdot w_2 = b$. Pero entonces

$$u \cdot v = u \cdot (w_2 - w_1) = u \cdot w_2 - u \cdot w_1 = b - b = 0$$

De este modo, $v = \overrightarrow{PQ}$ es ortogonal al vector normal u.

2.34. Encontrar una ecuación del hiperplano H en $\mathbf{R}^4$ que pasa por el punto $P(3, -2, 1, -4)$ y es normal al vector $u = (2, 5, -6, -2)$.

La ecuación de H es de la forma $2x + 5y - 6z - 2w = k$, puesto que es normal a u. Sustituyendo P en esta ecuación obtenemos $k = -2$. Por consiguiente, una ecuación de H es $2x + 5y - 6z - 2w = -2$.

2.35. Hallar una ecuación del plano H en $\mathbf{R}^3$ que contiene $P(1, -5, 2)$ y es paralelo al plano H' determinado por $3x - 7y + 4z = 5$.

H y H' son paralelos si y sólo si sus normales son paralelas o antiparalelas. Por tanto, una ecuación de H será de la forma $3x - 7y + 4z = k$. Sustituyendo $P(1, -5, 2)$ en la ecuación obtenemos $k = 46$. En consecuencia, la ecuación requerida es $3x - 7y + 4z = 46$.

2.36. Hallar una representación paramétrica de la recta en $\mathbf{R}^4$ que pasa por el punto $P(4, -2, 3, 1)$, en la dirección de $u = [2, 5, -7, 11]$.

La recta L en $\mathbf{R}^n$ que pasa por el punto $P(a_i)$, en la dirección del vector no nulo $u = [u_i]$ consta de los puntos $X = (x_i)$ que satisfacen la ecuación

$$X = P + tu \qquad \text{o} \qquad x_i = a_i + u_i t \qquad (\text{para } i = 1, 2, \ldots, n) \qquad [1]$$

donde el parámetro t toma todos los valores reales. Así obtenemos

$$\begin{cases} x = 4 + 2t \\ y = -2 + 5t \\ z = 3 - 7t \\ w = 1 + 11t \end{cases} \qquad \text{o} \qquad (4 + 2t,\ -2 + 5t,\ 3 - 7t,\ 1 + 11t)$$

2.37. Hallar una representación paramétrica de la recta en $\mathbf{R}^3$ que pasa por los puntos $P(5, 4, -3)$ y $Q(1, -3, 2)$.

Primero calculamos $u = \overrightarrow{PQ} = [1 - 5,\ -3 - 4,\ 2 - (-3)] = [-4,\ -7,\ 5]$. Entonces utilizamos el Problema 2.36 para obtener

$$x = 5 - 4t \qquad y = 4 - 7t \qquad z = -3 + 5t$$

2.38. Dar una representación no paramétrica para la recta del Problema 2.37.

Despejamos t de la ecuación de cada coordenada e igualamos los resultados llegando a

$$\frac{x - 5}{-4} = \frac{y - 4}{-7} = \frac{z + 3}{5}$$

o al par de ecuaciones lineales $7x - 4y = 19$ y $5x + 4z = 13$.

2.39. Encontrar una ecuación paramétrica de la recta en $\mathbf{R}^3$ perpendicular al plano $2x - 3y + 7z = 4$, que intersecta al mismo en el punto $P(6, 5, 1)$.

Como la recta es perpendicular al plano, debe estar en la dirección de su vector normal $u = [2, -3, 7]$. De donde

$$x = 6 + 2t \qquad y = 5 - 3t \qquad z = 1 + 7t$$

2.40. Considérese la siguiente curva C en $\mathbf{R}^4$, donde $0 \le t \le 4$:

$$F(t) = (t^2,\ 3t - 2,\ t^3,\ t^2 + 5)$$

Determinar el vector tangente unitario $\mathbf{T}$ para $t = 2$.

Tomamos la derivada de (cada componente de) $F(t)$ para obtener un vector V que sea tangente a la curva:

$$V(t) = \frac{dF(t)}{dt} = (2t,\ 3,\ 3t^2,\ 2t)$$

Ahora hallamos el valor de V para $t = 2$. Esto conduce a $V = (4, 3, 12, 4)$. Normalizamos V para obtener el vector unitario $\mathbf{T}$ tangente a la curva cuando $t = 2$. Tenemos

$$\|V\|^2 = 16 + 9 + 144 + 16 = 185 \qquad \text{o} \qquad \|V\| = \sqrt{185}$$

Así

$$\mathbf{T} = \left[\frac{4}{\sqrt{185}},\ \frac{3}{\sqrt{185}},\ \frac{12}{\sqrt{185}},\ \frac{4}{\sqrt{185}}\right]$$

2.41. Sea $\mathbf{T}(t)$ el vector unitario tangente a una curva C en $\mathbf{R}^n$. Demostrar que $d\mathbf{T}(t)/dt$ es ortogonal a $\mathbf{T}(t)$.

Tenemos $\mathbf{T}(t) \cdot \mathbf{T}(t) = 1$. Utilizando la regla de derivación del producto escalar, junto con $d(1)/dt = 0$, llegamos a

$$d[\mathbf{T}(t) \cdot \mathbf{T}(t)]/dt = \mathbf{T}(t) \cdot d\mathbf{T}(t)/dt + d\mathbf{T}(t)/dt \cdot \mathbf{T}(t) = 2\mathbf{T}(t) \cdot d\mathbf{T}(t)/dt = 0$$

De este modo, $d\mathbf{T}(t)/dt$ es ortogonal a $\mathbf{T}(t)$.

VECTORES ESPACIALES (EN $\mathbf{R}^3$). PLANOS, RECTAS, CURVAS Y SUPERFICIES EN $\mathbf{R}^3$

Las siguientes fórmulas se utilizarán en los Problemas 2.42-2.53.

La ecuación de un plano que pasa por el punto $P_0(x_0, y_0, z_0)$ con dirección normal $\mathbf{N} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ es

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \qquad [2.1]$$

La ecuación paramétrica de una recta L que pasa por el punto $P_0(x_0, y_0, z_0)$ en la dirección del vector $v = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ es

$$x = at + x_0 \qquad y = bt + y_0 \qquad z = ct + z_0$$

o, equivalentemente,

$$L(t) = (at + x_0)\mathbf{i} + (bt + y_0)\mathbf{j} + (ct + z_0)\mathbf{k} \qquad [2.2]$$

La ecuación de un vector $\mathbf{N}$ normal a una superficie $F(x, y, z) = 0$ es

$$\mathbf{N} = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k} \qquad [2.3]$$

2.42. Hallar la ecuación del plano con dirección normal $\mathbf{N} = 5\mathbf{i} - 6\mathbf{j} + 7\mathbf{k}$ que contiene el punto $P(3, 4, -2)$.

Sustituyendo P y $\mathbf{N}$ en la ecuación [2.1] llegamos a

$$5(x - 3) - 6(y - 4) + 7(z + 2) = 0 \qquad \text{o} \qquad 5x - 6y + 7z = -23$$

2.43. Hallar un vector $\mathbf{N}$ normal al plano $4x + 7y - 12z = 3$.

Los coeficientes de x, y, z dan una dirección normal, luego $\mathbf{N} = 4\mathbf{i} + 7\mathbf{j} - 12\mathbf{k}$. (Cualquier múltiplo de $\mathbf{N}$ es también normal al plano.)

2.44. Encontrar el plano H paralelo a $4x + 7y - 12z = 3$ que contiene el punto $P(2, 3, -1)$.

El plano dado y H tienen la misma dirección normal; es decir, $\mathbf{N} = 4\mathbf{i} + 7\mathbf{j} - 12\mathbf{k}$ es normal a H. Sustituyendo P y $\mathbf{N}$ en la ecuación a) se tiene:

$$4(x - 2) + 7(y - 3) - 12(z + 1) = 0 \qquad \text{o} \qquad 4x + 7y - 12z = 41$$

2.45. Sean H y K, respectivamente, los planos $x + 2y - 4z = 5$ y $2x - y + 3z = 7$. Hallar $\cos\theta$ siendo θ el ángulo entre ambos planos.

El ángulo θ entre los planos H y K es el ángulo entre la normal a H, **N**, y la normal a K, **N**′. Tenemos

$$\mathbf{N} = \mathbf{i} + 2\mathbf{j} - 4\mathbf{k} \qquad \text{y} \qquad \mathbf{N}' = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$$

Entonces

$$\mathbf{N} \cdot \mathbf{N}' = 2 - 2 - 12 = -12 \qquad \|\mathbf{N}\|^2 = 1 + 4 + 16 = 21 \qquad \|\mathbf{N}'\|^2 = 4 + 1 + 9 = 14$$

Así

$$\cos\theta = \frac{\mathbf{N} \cdot \mathbf{N}'}{\|\mathbf{N}\|\|\mathbf{N}'\|} = -\frac{12}{\sqrt{21}\sqrt{14}} = -\frac{12}{7\sqrt{6}}$$

2.46. Deducir la ecuación [2.1].

Sea $P(x, y, z)$ un punto arbitrario. El vector v de P_0 a P es

$$v = P - P_0 = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}$$

Como v es ortogonal a $\mathbf{N} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ (Fig. 2-12), llegamos a la fórmula requerida

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

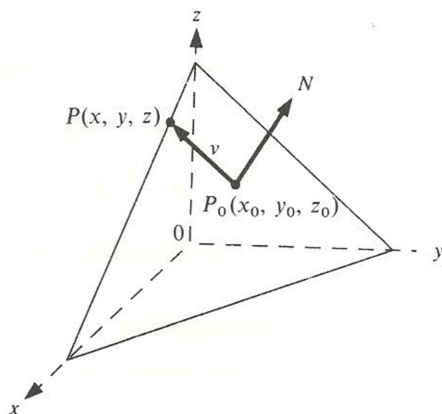


Figura 2-12.

2.47. Deducir la ecuación [2.2].

Sea $P(x, y, z)$ un punto arbitrario en la recta L. El vector w de P_0 a P es

$$w = P - P_0 = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k} \qquad [1]$$

Dado que w y v tienen la misma dirección (Fig. 2-13),

$$\mathbf{w} = t\mathbf{v} = t(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) = at\mathbf{i} + bt\mathbf{j} + ct\mathbf{k} \qquad [2]$$

Las ecuaciones [1] y [2] nos conducen al resultado buscado.

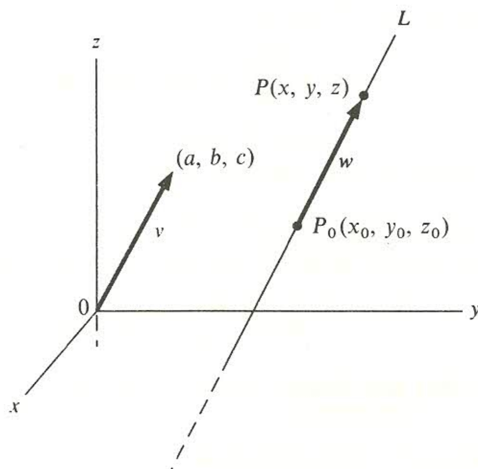


Figura 2-13.

2.48. Hallar la ecuación (paramétrica) de la recta L que pasa por:

a) El punto $P(3, 4, -2)$ en la dirección de $v = 5\mathbf{i} - \mathbf{j} + 3\mathbf{k}$.

b) Los puntos $P(1, 3, 2)$ y $Q(2, 5, -6)$.

a) Sustituimos en la ecuación [2.2] llegando a

$$L(t) = (5t + 3)\mathbf{i} + (-t + 4)\mathbf{j} + (3t - 2)\mathbf{k}$$

b) Determinamos, en primer lugar, el vector v de P a Q: $v = Q - P = \mathbf{i} + 2\mathbf{j} - 8\mathbf{k}$. A continuación utilizamos [2.2] con v y uno de los puntos dados, digamos P, obteniendo

$$L(t) = (t + 1)\mathbf{i} + (2t + 3)\mathbf{j} + (-8t + 2)\mathbf{k}$$

2.49. Sea H el plano $3x + 5y + 7z = 15$. Encontrar la ecuación de la recta L que, siendo perpendicular a H, contiene el punto $P(1, -2, 4)$.

Dado que L es perpendicular a H, debe tener la misma dirección que la normal $\mathbf{N} = 3\mathbf{i} + 5\mathbf{j} + 7\mathbf{k}$ a H. Así utilizando [2.2] con $\mathbf{N}$ y P se llega a

$$L(t) = (3t + 1)\mathbf{i} + (5t - 2)\mathbf{j} + (7t + 4)\mathbf{k}$$

2.50. Considérese un móvil B cuya posición en el instante t viene dada por $R(t) = t^3\mathbf{i} + 2t^2\mathbf{j} + 3t\mathbf{k}$. [Entonces $V(t) = dR(t)/dt$ denota la velocidad de B y $A(t) = dV(t)/dt$ su aceleración.]

a) Determinar la posición de B cuando $t = 1$.

b) Determinar la velocidad, v, de B cuando $t = 1$.

c) Determinar la rapidez, s, de B cuando $t = 1$.

d) Determinar la aceleración, a, de B cuando $t = 1$.

a) Sustituimos $t = 1$ en $R(t)$ y llegamos a $R(1) = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.

b) Tomamos la derivada de $R(t)$ obteniendo

$$V(t) = \frac{dR(t)}{dt} = 3t^2\mathbf{i} + 4t\mathbf{j} + 3\mathbf{k}$$

Sustituyendo $t = 1$ en $V(t)$, $v = V(1) = 3\mathbf{i} + 4\mathbf{j} + 3\mathbf{k}$.

c) La rapidez, s, es la magnitud de la velocidad v. De modo que

$$s^2 = \|v\|^2 = 9 + 16 + 9 = 34 \qquad \text{y, de aquí,} \qquad s = \sqrt{34}$$

d) Tomamos la derivada segunda de $R(t)$ o, en otras palabras, la derivada de $V(t)$ llegando a

$$A(t) = \frac{dV(t)}{dt} = 6t\mathbf{i} + 4\mathbf{j}$$

Sustituimos $t = 1$ en $A(t)$ para obtener $a = A(1) = 6\mathbf{i} + 4\mathbf{j}$.

2.51. Considérese la superficie $xy^2 + 2yz = 16$ en $\mathbf{R}^3$. Encontrar: *a*) el vector $\mathbf{N}(x, y, z)$ normal a la superficie, *b*) el plano H tangente a la superficie en el punto $P(1, 2, 3)$.

a) Calculamos las derivadas parciales F_x, F_y, F_z, donde $F(x, y, z) = xy^2 + 2yz - 16$. Tenemos

$$F_x = y^2 \qquad F_y = 2xy + 2z \qquad F_z = 2y$$

Así, según la ecuación [2.3], $\mathbf{N}(x, y, z) = y^2\mathbf{i} + (2xy + 2z)\mathbf{j} + 2y\mathbf{k}$.

b) La normal a la superficie en el punto P es

$$\mathbf{N}(P) = \mathbf{N}(1, 2, 3) = 4\mathbf{i} + 10\mathbf{j} + 4\mathbf{k}$$

Siendo así, $\mathbf{N} = 2\mathbf{i} + 5\mathbf{j} + 2\mathbf{k}$ es también un vector normal en P. Sustituyendo P y $\mathbf{N}$ en la ecuación [2.1] se obtiene

$$2(x - 1) + 5(y - 2) + 2(z - 3) = 0 \qquad \text{o} \qquad 2x + 5y + 2z = 18$$

2.52. Considérese el elipsoide $x^2 + 2y^2 + 3z^2 = 15$. Determinar el plano H, tangente en el punto $P(2, 2, 1)$.

Primero hallamos el vector normal (a partir de la ecuación [2.3])

$$\mathbf{N}(x, y, z) = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k} = 2x\mathbf{i} + 4y\mathbf{j} + 6z\mathbf{k}$$

Evaluamos el vector normal $\mathbf{N}(x, y, z)$ en el punto P llegando a

$$\mathbf{N}(P) = \mathbf{N}(2, 2, 1) = 4\mathbf{i} + 8\mathbf{j} + 6\mathbf{k}$$

Así $\mathbf{N} = 2\mathbf{i} + 4\mathbf{j} + 3\mathbf{k}$ es normal al elipsoide en P. Sustituyendo P y $\mathbf{N}$ en [2.1] obtendremos la ecuación de H:

$$2(x - 2) + 4(y - 2) + 3(z - 1) = 0 \qquad \text{o} \qquad 2x + 4y + 3z = 15$$

2.53. Considérese la función $f(x, y) = x^2 + y^2$ para la que $z = x^2 + y^2$ representa una superficie S en $\mathbf{R}^3$. Determinar: *a*) el vector $\mathbf{N}$ normal a la superficie para $x = 2$, $y = 3$, *b*) el plano H tangente a la superficie S para $x = 2$, $y = 3$.

a) Utilicemos el hecho de que, si $F(x, y, z) = f(x, y) - z$, tenemos $F_x = f_x$, $F_y = f_y$ y $F_z = -1$. Entonces

$$\mathbf{N} = (f_x, f_y, -1) = 2x\mathbf{i} + 2y\mathbf{j} - \mathbf{k} = 4\mathbf{i} + 6\mathbf{j} - \mathbf{k}$$

b) Si $x = 2$ e $y = 3$, entonces $z = 4 + 9 = 13$; por consiguiente, $P(2, 3, 13)$ es el punto en la superficie S. Sustituimos P y $\mathbf{N} = 4\mathbf{i} + 6\mathbf{j} - \mathbf{k}$ en la ecuación [2.1] para obtener H.

$$4(x - 2) + 6(y - 3) - (z - 13) = 0 \quad \text{o} \quad 4x + 6y - z = 13$$

PRODUCTO VECTORIAL

El producto vectorial está definido únicamente para vectores en $\mathbf{R}^3$.

2.54. Calcular $u \times v$, donde *a*) $u = (1, 2, 3)$ y $v = (4, 5, 6)$, *b*) $u = (7, 3, 1)$ y $v = (1, 1, 1)$, *c*) $u = (-4, 12, 2)$ y $v = (6, -18, 3)$.

El producto vectorial de $u = (a_1, a_2, a_3)$ y $v = (b_1, b_2, b_3)$ puede obtenerse como sigue. Situamos el vector $v = (b_1, b_2, b_3)$ *bajo* el $u = (a_1, a_2, a_3)$ construyendo la matriz

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

Entonces

$$u \times v = \left(\begin{vmatrix} \boxed{a_1} & a_2 & a_3 \\ \boxed{b_1} & b_2 & b_3 \end{vmatrix}, -\begin{vmatrix} a_1 & \boxed{a_2} & a_3 \\ b_1 & \boxed{b_2} & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 & \boxed{a_3} \\ b_1 & b_2 & \boxed{b_3} \end{vmatrix} \right)$$

Esto es, tapamos la primera columna de la matriz y tomamos el determinante para obtener la primera componente de $u \times v$; tapamos la segunda columna y tomamos el negativo del determinante para obtener la segunda componente, y tapamos la tercera columna y tomamos el determinante para obtener la tercera componente.

a) $u \times v = \left(\begin{vmatrix} \boxed{1} & 2 & 3 \\ \boxed{4} & 5 & 6 \end{vmatrix}, -\begin{vmatrix} 1 & \boxed{2} & 3 \\ 4 & \boxed{5} & 6 \end{vmatrix}, \begin{vmatrix} 1 & 2 & \boxed{3} \\ 4 & 5 & \boxed{6} \end{vmatrix} \right) = (12 - 15, 12 - 6, 5 - 8) = (-3, 6, -3).$

b) $u \times v = \left(\begin{vmatrix} \boxed{7} & 3 & 1 \\ \boxed{1} & 1 & 1 \end{vmatrix}, -\begin{vmatrix} 7 & \boxed{3} & 1 \\ 1 & \boxed{1} & 1 \end{vmatrix}, \begin{vmatrix} 7 & 3 & \boxed{1} \\ 1 & 1 & \boxed{1} \end{vmatrix} \right) = (3 - 1, 1 - 7, 7 - 3) = (2, -6, 4).$

c) $u \times v = \left(\begin{vmatrix} \boxed{-4} & 12 & 2 \\ \boxed{6} & -18 & 3 \end{vmatrix}, -\begin{vmatrix} -4 & \boxed{12} & 2 \\ 6 & \boxed{-18} & 3 \end{vmatrix}, \begin{vmatrix} -4 & 12 & \boxed{2} \\ 6 & -18 & \boxed{-3} \end{vmatrix} \right) =$

$= (36 + 36, 12 + 12, 72 - 72) = (72, 24, 0).$

2.55. Considérense los vectores $u = 2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$, $v = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, $w = \mathbf{i} + 5\mathbf{j} + 3\mathbf{k}$. Encontrar: *a*) $u \times v$, *b*) $u \times w$, *c*) $v \times w$.

Utilizamos

$$v_1 \times v_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}$$

donde $v_1 = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ y $v_2 = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$.

a) $u \times v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -3 & 4 \\ 3 & 1 & -2 \end{vmatrix} = (6 - 4)\mathbf{i} + (12 + 4)\mathbf{j} + (2 + 9)\mathbf{k} = 2\mathbf{i} + 16\mathbf{j} + 11\mathbf{k}.$

(*Nota:* Obsérvese que la componente **j** se obtiene tomando el determinante cambiado de signo. Véase el Problema 2.54.)

b) $u \times w = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -3 & 4 \\ 1 & 5 & 3 \end{vmatrix} = (-9 - 20)\mathbf{i} + (4 - 6)\mathbf{j} + (10 + 3)\mathbf{k} = -29\mathbf{i} - 2\mathbf{j} + 13\mathbf{k}.$

c) $v \times w = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & -2 \\ 1 & 5 & 3 \end{vmatrix} = (3 + 10)\mathbf{i} + (-2 - 9)\mathbf{j} + (15 - 1)\mathbf{k} = 13\mathbf{i} - 11\mathbf{j} + 14\mathbf{k}.$

2.56. Demostrar el Teorema 2.5 i): El vector $u \times v$ es ortogonal a u y a v.

Supongamos $u = (a_1, a_2, a_3)$ y $v = (b_1, b_2, b_3)$. Entonces

$$u \cdot (u \times v) = a_1(a_2 b_3 - a_3 b_2) + a_2(a_3 b_1 - a_1 b_3) + a_3(a_1 b_2 - a_2 b_1) =$$
$$= a_1 a_2 b_3 - a_1 a_3 b_2 + a_2 a_3 b_1 - a_1 a_2 b_3 + a_1 a_3 b_2 - a_2 a_3 b_1 = 0$$

De este modo, $u \times v$ es ortogonal a u. Análogamente se demuestra que es ortogonal a v.

2.57. Encontrar un vector unitario u ortogonal a $v = (1, 3, 4)$ y a $w = (2, -6, 5)$.

Comenzamos calculando $v \times w$ que es ortogonal a ambos vectores. La matriz $\begin{pmatrix} 1 & 3 & 4 \\ 2 & -6 & -5 \end{pmatrix}$ conduce a

$$v \times w = (-15 + 24, 8 + 5, -6 - 6) = (9, 13, -12)$$

Ahora normalizamos $u \times w$ obteniendo $u = (9/\sqrt{394}, 13/\sqrt{394}, -12/\sqrt{394})$.

2.58. Demostrar la *identidad de Lagrange*: $\|u \times v\|^2 = (u \cdot u)(v \cdot v) - (u \cdot v)^2$.

Si $u = (a_1, a_2, a_3)$ y $v = (b_1, b_2, b_3)$, entonces

$$\|u \times v\|^2 = (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2 \quad [1]$$

$$(u \cdot u)(v \cdot v) - (u \cdot v)^2 = (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2 \quad [2]$$

Desarrollando la parte derecha de [1] y [2] se llega a la identidad.

NUMEROS COMPLEJOS

2.59. Supóngase que $z = 5 + 3i$ y $w = 2 - 4i$. Calcular: *a*) $z + w$, *b*) $z - w$, *c*) zw.

Utilizamos las reglas algebraicas ordinarias, junto con la igualdad $i^2 = -1$, para llegar a un resultado en la forma convencional $a + bi$.

a) $z + w = (5 + 3i) + (2 - 4i) = 7 - i.$

b) $z - w = (5 + 3i) - (2 - 4i) = 5 + 3i - 2 + 4i = 3 + 7i.$

c) $zw = (5 + 3i)(2 - 4i) = 10 - 14i - 12i^2 = 10 - 14i + 12 = 22 - 14i.$

2.60. Simplificar: *a)* $(5+3i)(2-7i)$, *b)* $(4-3i)^2$, *c)* $(1+2i)^3$.

a) $(5+3i)(2-7i) = 10+6i-35i-21i^2 = 31-29i$.

b) $(4-3i)^2 = 16-24i+9i^2 = 7-24i$.

c) $(1+2i)^3 = 1+6i+12i^2+8i^3 = 1+6i-12-8i = -11-2i$.

2.61. Simplificar: *a)* i^0, i^3, i^4, *b)* i^5, i^6, i^7, i^8, *c)* i^{39}, i^{174}, i^{252}, i^{317}.

a) $i^0 = 1$, $i^3 = i^2(i) = (-1)(i) = -i$, $i^4 = (i^2)(i^2) = (-1)(-1) = 1$.

b) $i^5 = (i^4)(i) = (1)(i) = i$, $i^6 = (i^4)(i^2) = (1)(i^2) = i^2 = -1$, $i^7 = i^3 = -i$, $i^8 = i^4 = 1$.

c) Usando $i^4 = 1$ e $i^n = i^{4q+r} = (i^4)^q i^r = 1^q i^r = i^r$ dividimos el exponente n por 4 obteniendo el resto r:

$$i^{39} = i^{4(9)+3} = (i^4)^9 i^3 = 1^9 i^3 = i^3 = -i \qquad i^{174} = i^2 = -1 \qquad i^{252} = i^0 = 1 \qquad i^{317} = i^1 = i$$

2.62. Determinar el complejo conjugado de cada uno de los siguientes números:

a) $6+4i$, $7-5i$, $4+i$, $-3-i$; *b)* 6, -3, $4i$, $-9i$.

a) $\overline{6+4i} = 6-4i$, $\overline{7-5i} = 7+5i$, $\overline{4+i} = 4-i$, $\overline{-3-i} = -3+i$.

b) $\bar{6} = 6$, $\overline{-3} = -3$, $\overline{4i} = -4i$, $\overline{-9i} = 9i$.

(Nótese que el conjugado de un número real es el número original, pero el conjugado de un número imaginario puro es el opuesto del número original.)

2.63. Hallar $z\bar{z}$ y $|z|$, donde $z = 3+4i$.

Para $z = a+bi$ utilizamos $z\bar{z} = a^2+b^2$ y $z = \sqrt{z\bar{z}} = \sqrt{a^2+b^2}$.

$$z\bar{z} = 9+16 = 25 \qquad |z| = \sqrt{25} = 5$$

2.64. Simplificar $\dfrac{2-7i}{5+3i}$.

Para simplificar una fracción z/w de números complejos, multiplicamos tanto el numerador como el denominador por $\bar{w}$, el conjugado del denominador.

$$\frac{2-7i}{5+3i} = \frac{(2-7i)(5-3i)}{(5+3i)(5-3i)} = \frac{-11-41i}{34} = -\frac{11}{34} - \frac{41}{34}i$$

2.65. Demostrar las siguientes afirmaciones: Para dos números complejos cualesquiera $z, w \in \mathbf{C}$, i) $\overline{z+w} = \bar{z}+\bar{w}$, ii) $\overline{zw} = \bar{z}\bar{w}$, iii) $\bar{\bar{z}} = z$.

Supongamos $z = a+bi$ y $w = c+di$, donde $a, b, c, d \in \mathbf{R}$.

i) $$\begin{aligned}\overline{z+w} &= \overline{(a+bi)+(c+di)} = \overline{(a+c)+(b+d)i} = \\ &= (a+c)-(b+d)i = a+c-bi-di = \\ &= (a-bi)+(c-di) = \bar{z}+\bar{w}.\end{aligned}$$

ii) $$\begin{aligned}\overline{zw} &= \overline{(a+bi)(c+di)} = \overline{(ac-bd)+(ad+bc)i} = \\ &= (ac-bd)-(ad+bc)i = (a-bi)(c-di) = \bar{z}\bar{w}\end{aligned}$$

iii) $\bar{\bar{z}} = \overline{\overline{a+bi}} = \overline{a-bi} = a-(-b)i = a+bi = z$.

2.66. Demostrar la siguiente afirmación: Para dos números complejos cualesquiera z, $w \in \mathbf{C}$, $|zw| = |z|\,|w|$.

Supongamos $z = a + bi$ y $w = c + di$, donde $a, b, c, d \in \mathbf{R}$. Entonces

$$|z|^2 = a^2 + b^2 \qquad |w|^2 = c^2 + d^2 \qquad \text{y} \qquad zw = (ac - bd) + (ad + bc)i$$

Así

$$|zw|^2 = (ac - bd)^2 + (ad + bc)^2 =$$
$$= a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2 =$$
$$= a^2(c^2 + d^2) + b^2(c^2 + d^2) = (a^2 + b^2)(c^2 + d^2) = |z|^2|w|^2$$

La raíz cuadrada de ambos miembros proporciona el resultado deseado.

2.67. Demostrar la siguiente afirmación: Para dos números complejos cualesquiera z, $w \in \mathbf{C}$, $|z + w| \le |z| + |w|$.

Supongamos $z = a + bi$ y $w = c + di$, con $a, b, c, d \in \mathbf{R}$. Consideremos los vectores $u = (a, b)$ y $v = (c, d)$ en $\mathbf{R}^2$. Nótese que

$$|z| = \sqrt{a^2 + b^2} = \|u\| \qquad |w| = \sqrt{c^2 + d^2} = \|v\|$$

y

$$|z + w| = |(a + c) + (b + d)i| = \sqrt{(a + c)^2 + (b + d)^2} = \|(a + c, b + d)\| = \|u + v\|$$

Según la desigualdad de Minkowski (Problema 2.23), $\|u + v\| \le \|u\| + \|v\|$ y por tanto

$$|z + w| = \|u + v\| \le \|u\| + \|v\| = |z| + |w|$$

2.68. Hallar los productos escalares $u \cdot v$ y $v \cdot u$, donde: *a*) $u = (1 - 2i, 3 + i)$, $v = (4 + 2i, 5 - 6i)$; *b*) $u = (3 - 2i, 4i, 1 + 6i)$, $v = (5 + i, 2 - 3i, 7 + 2i)$.

Recuérdese que los conjugados de las componentes del segundo vector aparecen en el producto escalar

$$(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n$$

a) $u \cdot v = (1 - 2i)\overline{(4 + 2i)} + (3 + i)\overline{(5 - 6i)} =$

$= (1 - 2i)(4 - 2i) + (3 + i)(5 + 6i) = -10i + 9 + 23i = 9 + 13i$

$v \cdot u = (4 + 2i)\overline{(1 - 2i)} + (5 - 6i)\overline{(3 + i)} =$

$= (4 + 2i)(1 + 2i) + (5 - 6i)(3 - i) = 10i + 9 - 23i = 9 - 13i$

b) $u \cdot v = (3 - 2i)\overline{(5 + i)} + (4i)\overline{(2 - 3i)} + (1 + 6i)\overline{(7 + 2i)} =$

$= (3 - 2i)(5 - i) + (4i)(2 + 3i) + (1 + 6i)(7 - 2i) = 20 + 35i$

$v \cdot u = (5 + i)\overline{(3 - 2i)} + (2 - 3i)\overline{(4i)} + (7 + 2i)\overline{(1 + 6i)} =$

$= (5 + i)(3 + 2i) + (2 - 3i)(-4i) + (7 + 2i)(1 - 6i) = 20 - 35i$

En ambos casos, $v \cdot u = \overline{u \cdot v}$. Esto es cierto en general, como se verá en el Problema 2.70.

2.69. Sean $u = (7 - 2i, 2 + 5i)$ y $v = (1 + i, -3 - 6i)$. Calcular:

a) $u + v$; *b)* $2iu$; *c)* $(3 - i)v$; *d)* $u \cdot v$; *e)* $\|u\|$ y $\|v\|$.

a) $u + v = (7 - 2i + 1 + i, 2 + 5i - 3 - 6i) = (8 - i, -1 - i)$.

b) $2iu = (14i - 4i^2, 4i + 10i^2) = (4 + 14i, -10 + 4i)$.

c) $(3 - i)v = (3 + 3i - i - i^2, -9 - 18i + 3i + 6i^2) = (4 + 2i, -15 - 15i)$.

d) $u \cdot v = (7 - 2i)\overline{(1 + i)} + (2 + 5i)\overline{(-3 - 6i)} =$
$= (7 - 2i)(1 - i) + (2 + 5i)(-3 + 6i) = 5 - 9i - 36 - 3i = -31 - 12i$.

e) $\|u\| = \sqrt{7^2 + (-2)^2 + 2^2 + 5^2} = \sqrt{82}$, $\|v\| = \sqrt{1^2 + 1^2 + (-3)^2 + (-6)^2} = \sqrt{47}$.

2.70. Demostrar las siguientes afirmaciones: Para dos vectores cualesquiera $u, v \in \mathbf{C}^n$ y cualquier escalar $z \in \mathbf{C}$, i) $u \cdot v = \overline{v \cdot u}$, ii) $(zu) \cdot v = z(u \cdot v)$, iii) $u \cdot (zv) = \bar{z}(u \cdot v)$.

Supongamos $u = (z_1, z_2, \ldots, z_n)$ y $v = (w_1, w_2, \ldots, w_n)$.

i) Usando las propiedades de la conjugación,

$$\overline{v \cdot u} = \overline{w_1 \bar{z}_1 + w_2 \bar{z}_2 + \cdots + w_n \bar{z}_n} = \overline{w_1 \bar{z}_1} + \overline{w_2 \bar{z}_2} + \cdots + \overline{w_n \bar{z}_n} =$$
$$= \bar{w}_1 z_1 + \bar{w}_2 z_2 + \cdots + \bar{w}_n z_n = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n = u \cdot v$$

ii) Como $zu = (zz_1, zz_2, \ldots, zz_n)$,

$$(zu) \cdot v = zz_1 \bar{w}_1 + zz_2 \bar{w}_2 + \cdots + zz_n \bar{w}_n = z(z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n) = z(u \cdot v)$$

(Compárese con el Teorema 2.2, referente a vectores en $\mathbf{R}^n$.)

iii) **Método 1.** Dado que $zv = (zw_1, zw_2, \ldots, zw_n)$,

$$u \cdot (zv) = z_1 \overline{zw_1} + z_2 \overline{zw_2} + \cdots + z_n \overline{zw_n} = z_1 \bar{z} \bar{w}_1 + z_2 \bar{z} \bar{w}_2 + \cdots + z_n \bar{z} \bar{w}_n =$$
$$= \bar{z}(z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n) = \bar{z}(u \cdot v)$$

Método 2. Utilizando i) y ii),

$$u \cdot (zv) = \overline{(zv) \cdot u} = \overline{z(v \cdot u)} = \bar{z}\overline{(v \cdot u)} = \bar{z}(u \cdot v)$$

PROBLEMAS SUPLEMENTARIOS

VECTORES EN $\mathbf{R}^n$

2.71. Sean $u = (2, -1, 0, -3)$, $v = (1, -1, -1, 3)$, $w = (1, 3, -2, 2)$. Hallar: *a)* $2u - 3v$; *b)* $5u - 3v - 4w$; *c)* $-u + 2v - 2w$; *d)* $u \cdot v$, $u \cdot w$ y $v \cdot w$; *e)* $\|u\|$, $\|v\|$ y $\|w\|$.

2.72. Determinar x e y si: *a)* $x(3, 2) = 2(y, -1)$; *b)* $x(2, y) = y(1, -2)$.

2.73. Hallar $d(u, v)$ y $\text{proy}\,(u, v)$ cuando: *a)* $u = (1, -3)$, $v = (4, 1)$; *b)* $u = (2, -1, 0, 1)$, $v = (1, -1, 1, 2)$.

COMBINACIONES LINEALES. INDEPENDENCIA LINEAL

2.74. Sean

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad u_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Expresamos v como combinación lineal de u_1, u_2, u_3, donde

$$a)\quad v = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}, \qquad b)\quad v = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \qquad c)\quad v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

2.75. Determinar si los siguientes vectores u, v, w son linealmente independientes y, en caso de no serlo, expresar uno de ellos como combinación lineal de los otros.

a) $u = (1, 0, 1)$, $v = (1, 2, 3)$, $w = (3, 2, 5)$.

b) $u = (1, 0, 1)$, $v = (1, 1, 1)$, $w = (0, 1, 1)$.

c) $u = (1, 2)$, $v = (1, -1)$, $w = (2, 5)$.

d) $u = (1, 0, 0, 1)$, $v = (0, 1, 2, 1)$, $w = (1, 2, 3, 4)$.

e) $u = (1, 0, 0, 1)$, $v = (0, 1, 2, 1)$, $w = (1, 2, 4, 3)$.

VECTORES LOCALIZADOS. HIPERPLANOS, RECTAS, CURVAS

2.76. Encontrar el vector (localizado) v de *a)* $P(2, 3, -7)$ a $Q(1, -6, -5)$; *b)* $P(1, -8, -4, 6)$ a $Q(3, -5, 2, -4)$.

2.77. Hallar una ecuación del hiperplano en $\mathbf{R}^3$ que:

a) Pasa por $(2, -7, 1)$ y es normal a $(3, 1, -11)$.

b) Contiene $(1, -2, 2)$, $(0, 1, 3)$ y $(0, 2, -1)$.

c) Contiene $(1, -5, 2)$ y es paralelo a $3x - 7y + 4z = 5$.

2.78. Encontrar una representación paramétrica de la recta que:

a) Pasa por $(7, -1, 8)$ en la dirección de $(1, 3, -5)$.

b) Pasa por $(1, 9, -4, 5)$ y $(2, -3, 0, 4)$.

c) Pasa por $(4, -1, 9)$ y es perpendicular al plano $3x - 2y + z = 18$.

VECTORES ESPACIALES (EN $\mathbf{R}^3$). PLANOS, RECTAS, CURVAS Y SUPERFICIES EN $\mathbf{R}^3$

2.79. Hallar la ecuación del plano H:

a) Con normal $\mathbf{N} = 3\mathbf{i} - 4\mathbf{j} + 5\mathbf{k}$, que contiene el punto $P(1, 2, -3)$.

b) Paralelo a $4x - 3y + 2z = 11$, que contiene el punto $P(1, 2, -3)$.

2.80. Encontrar un vector unitario u que sea normal al plano:

a) $3x - 4y - 12z = 11$; *b*) $2x - y - 2z = 7$.

2.81. Hallar cos θ, siendo θ el ángulo entre los planos:

a) $3x - 2y - 4z = 5$ y $x + 2y - 6z = 4$.

b) $2x + 5y - 4z = 1$ y $4x + 3y + 2z = 1$.

2.82. Determinar la ecuación (paramétrica) de la recta L:

a) Que pasa por el punto $P(2, 5, -3)$, en la dirección de $v = 4\mathbf{i} - 5\mathbf{j} + 7\mathbf{k}$.

b) Que pasa por los puntos $P(1, 2, -4)$ y $Q(3, -7, 2)$.

c) Perpendicular al plano $2x - 3y + 7z = 4$ y que contiene el punto $P(1, -5, 7)$.

2.83. Considérese la siguiente curva, donde $0 \le t \le 5$:

$$F(t) = t^3\mathbf{i} - t^2\mathbf{j} + (2t - 3)\mathbf{k}$$

a) Hallar $F(t)$ cuando $t = 2$.

b) Determinar los extremos de la curva.

c) Encontrar el vector unitario $\mathbf{T}$ tangente a la curva cuando $t = 2$.

2.84. Considérese la curva $F(t) = (\cos t)\mathbf{i} + (\operatorname{sen} t)\mathbf{j} + t\mathbf{k}$.

a) Determinar el vector unitario $\mathbf{T}(t)$ tangente a la curva.

b) Hallar el vector unitario $\mathbf{N}(t)$ normal a la curva normalizando $U(t) = d\mathbf{T}(t)/dt$.

c) Encontrar el vector unitario $\mathbf{B}(t)$ binormal a la curva utilizando $\mathbf{B} = \mathbf{T} \times \mathbf{N}$.

2.85. Considérese un móvil B cuya posición en el instante t viene dada por $\mathbf{R}(t) = t^2\mathbf{i} + t^3\mathbf{j} + 2t\mathbf{k}$. [Entonces $V(t) = dR(t)/dt$ denota la velocidad de B y $A(t) = dV(t)/dt$ su aceleración.]

a) Determinar la posición de B cuando $t = 1$.

b) Determinar la velocidad, v, de B cuando $t = 1$.

c) Determinar la rapidez, s, de B cuando $t = 1$.

d) Determinar la aceleración, a, de B cuando $t = 1$.

2.86. Hallar el vector normal $\mathbf{N}$ y el plano tangente H a la superficie en el punto dado:

a) Superficie $x^2y + 3yz = 20$ y punto $P(1, 3, 2)$.

b) Superficie $x^2 + 3y^2 - 5z^2 = 16$ y punto $P(3, -2, 1)$.

2.87. Dada $z = f(x, y) = x^2 + 2xy$, encontrar el vector normal $\mathbf{N}$ y el plano tangente H para $x = 3$, $y = 1$.

PRODUCTO VECTORIAL

El producto vectorial sólo está definido para vectores en $\mathbf{R}^3$.

2.88. Dados $u = 3\mathbf{i} - 4\mathbf{j} + 2\mathbf{k}$, $v = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$, $w = 4\mathbf{i} + 7\mathbf{j} + 2\mathbf{k}$. Hallar: *a*) $u \times v$, *b*) $u \times w$, *c*) $v \times w$, *d*) $v \times u$.

2.89. Hallar un vector unitario w ortogonal a *a*) $u = (1, 2, 3)$ y $v = (1, -1, 2)$; *b*) $u = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ y $v = 4\mathbf{i} - 2\mathbf{j} - \mathbf{k}$.

2.90. Demostrar las siguientes propiedades del producto vectorial:

a) $u \times v = -(v \times u)$

b) $u \times u = 0$ para todo vector u

c) $(ku) \times v = k(u \times v) = u \times (kv)$

d) $u \times (v + w) = (u \times v) + (u \times w)$

e) $(v + w) \times u = (v \times u) + (w \times u)$

f) $(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$

NUMEROS COMPLEJOS

2.91. Simplificar: *a*) $(4 - 7i)(9 + 2i)$; *b*) $(3 - 5i)^2$; *c*) $\dfrac{1}{4 - 7i}$; *d*) $\dfrac{9 + 2i}{3 - 5i}$; *e*) $(1 - i)^3$.

2.92. Simplificar: *a*) $\dfrac{1}{2i}$; *b*) $\dfrac{2 + 3i}{7 - 3i}$; *c*) i^{15}, i^{25}, i^{34}; *d*) $\left(\dfrac{1}{3 - i}\right)^2$.

2.93. Sean $z = 2 - 5i$ y $w = 7 + 3i$. Calcular: *a*) $z + w$; *b*) zw; *c*) z/w; *d*) $\bar{z}$, $\bar{w}$; *e*) $|z|$, $|w|$.

2.94. Sean $z = 2 + i$ y $w = 6 - 5i$. Calcular: *a*) z/w; *b*) $\bar{z}$, $\bar{w}$; *c*) $|z|$, $|w|$.

2.95. Demostrar que *a*) $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$; *b*) $\operatorname{Im} z = (z - \bar{z})/2i$.

2.96. Demostrar que $zw = 0$ implica $z = 0$ o $w = 0$.

VECTORES EN $\mathbf{C}^n$

2.97. Demostrar las siguientes afirmaciones: Para vectores arbitrarios $u, v, w \in \mathbf{C}^n$:

i) $(u + v) \cdot w = u \cdot w + v \cdot w$; ii) $w \cdot (u + v) = w \cdot u + w \cdot v$.

2.98. Probar que la norma en $\mathbf{C}^n$ satisface:

[N_1] Para cualquier vector u, $\|u\| \geq 0$; y $\|u\| = 0$ si y sólo si $u = 0$.

[N_2] Para cualquier vector u y cualquier número complejo z, $\|zu\| = |z|\,\|u\|$.

[N_3] Para vectores cualesquiera u y v, $\|u + v\| \leq \|u\| + \|v\|$.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

2.71. *a*) $2u - 3v = (1, 1, 3, -15)$

b) $5u - 3v - 4w = (3, -14, 11, -32)$

c) $-u + 2v - 2w = (-2, -7, 2, 5)$

d) $u \cdot v = -6$, $u \cdot w = -7$, $v \cdot w = 6$

e) $\|u\| = \sqrt{14}$, $\|v\| = 2\sqrt{3}$, $\|w\| = 3\sqrt{2}$

2.72. *a*) $x = -1$, $y = -\frac{3}{2}$; *b*) $x = 0$, $y = 0$ o $x = -2$, $y = -4$.

2.73. *a*) $d = 5$, $\operatorname{proy}(u, v) = (\frac{4}{17}, \frac{1}{17})$; *b*) $d = \sqrt{11}$, $\operatorname{proy}(u, v) = (\frac{5}{7}, -\frac{5}{7}, \frac{5}{7}\ \frac{10}{7})$.

2.74. *a*) $v = (\frac{9}{2})u_1 - 5u_2 + 3u_3$.
b) $v = u_2 + 2u_3$.
c) $v = [(a - 2b + c)/2]u_1 + (a + b - c)u_2 + [(c - a)/2]u_3$.

2.75. *a*) dependientes; *b*) independientes; *c*) dependientes; *d*) independientes; *e*) dependientes.

2.76. *a*) $v = (-1, -9, 2)$; *b*) $v = (2, 3, 6, -10)$.

2.77. *a*) $3x + y - 11z = -12$; *b*) $13x + 4y + z = 7$; *c*) $3x - 7y + 4z = 46$.

2.78. *a*) $x = 7 + t, y = -1 + 3t, z = 8 - 5t$.
b) $x_1 = 1 + t, x_2 = 9 - 12t, x_3 = -4 + 4t, x_4 = 5 - t$.
c) $x = 4 + 3t, y = -1 - 2t, z = 9 + t$.

2.79. *a*) $3x - 4y + 5z = -20$; *b*) $4x + 3y - 2z = 16$.

2.80. *a*) $u = (3\mathbf{i} - 4\mathbf{j} - 12\mathbf{k})/13$; *b*) $u = (2\mathbf{i} - \mathbf{j} - 2\mathbf{k})/3$.

2.81. *a*) $23/(\sqrt{29}\sqrt{41})$; *b*) $15/(\sqrt{45}\sqrt{29})$.

2.82. *a*) $x = 2 + 4t, y = 5 - 5t, z = -3 + 7t$.
b) $x = 1 + 2t, y = 2 - 9t, z = -4 + 6t$.
c) $x = 1 + 2t, y = -5 - 3t, z = 7 + 7t$.

2.83. *a*) $8\mathbf{i} - 4\mathbf{j} + \mathbf{k}$; *b*) $-3\mathbf{k}$ y $125\mathbf{i} - 25\mathbf{j} + 7\mathbf{k}$; *c*) $\mathbf{T} = (6\mathbf{i} - 2\mathbf{j} + \mathbf{k})/\sqrt{41}$.

2.84. *a*) $\mathbf{T}(t) = (-\text{sen } t)\mathbf{i} + (\cos t)\mathbf{j} + \mathbf{k})/\sqrt{2}$.
b) $\mathbf{N}(t) = (-\cos t)\mathbf{i} - (\text{sen } t)\mathbf{j}$.
c) $\mathbf{B}(t) = (\text{sen } t)\mathbf{i} - (\cos t)\mathbf{j} + \mathbf{k})/\sqrt{2}$.

2.85. *a*) $\mathbf{i} + \mathbf{j} + 2\mathbf{k}$; *b*) $2\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$; *c*) $\sqrt{17}$; *d*) $2\mathbf{i} + 6\mathbf{j}$.

2.86. *a*) $\mathbf{N} = 6\mathbf{i} + 7\mathbf{j} + 9\mathbf{k}$, $6x + 7y + 9z = 45$.
b) $\mathbf{N} = 6\mathbf{i} - 12\mathbf{j} - 10\mathbf{k}$, $3x - 6y - 5z = 16$.

2.87. $\mathbf{N} = 8\mathbf{i} + 6\mathbf{j} - \mathbf{k}$, $8x + 6y - z = 15$.

2.88. *a*) $2\mathbf{i} + 13\mathbf{j} + 23\mathbf{k}$ *c*) $31\mathbf{i} - 16\mathbf{j} - 6\mathbf{k}$
b) $-22\mathbf{i} + 2\mathbf{j} + 37\mathbf{k}$ *d*) $-2\mathbf{i} - 13\mathbf{j} - 23\mathbf{k}$

2.89. *a*) $(7, 1, -3)/\sqrt{59}$; *b*) $(5\mathbf{i} + 11\mathbf{j} - 2\mathbf{k})/\sqrt{150}$.

2.91. *a*) $50 - 55i$; *b*) $-16 - 30i$; *c*) $(4 + 7i)/65$; *d*) $(1 + 3i)/2$; *e*) $-2 - 2i$.

2.92. *a*) $-\frac{1}{2}i$; *b*) $(5 + 27i)/58$; *c*) $-i, i, -1$; *d*) $(4 + 3i)/50$.

2.93. *a*) $z + w = 9 - 2i$ *d*) $\bar{z} = 2 + 5i$, $\bar{w} = 7 - 3i$

b) $zw = 29 - 29i$ *e*) $|z| = \sqrt{29}$, $|w| = \sqrt{58}$

c) $z/w = (-1 - 41i)/58$

2.94. *a*) $z/w = (7 + 16i)/61$; *b*) $\bar{z} = 2 - i$, $\bar{w} = 6 + 5i$; *c*) $|z| = \sqrt{5}$, $|w| = \sqrt{61}$.

2.96. Si $zw = 0$, entonces $|zw| = |z|\,|w| = |0| = 0$. De aquí $|z| = 0$ o $|w| = 0$; y por tanto $z = 0$ o $w = 0$.

CAPITULO 3

Matrices

3.1. INTRODUCCION

Las matrices ya se trataron en el Capítulo 1, y sus elementos se relacionaron con los coeficientes de sistemas de ecuaciones lineales. Aquí introduciremos nuevamente esas matrices y estudiaremos ciertas operaciones algebraicas definidas sobre ellas. El material aquí expuesto está destinado principalmente al cálculo. No obstante, tal y como ocurría con las ecuaciones lineales, el tratamiento abstracto presentado posteriormente nos ayudará a profundizar en la estructura de las matrices.

Las entradas de nuestras matrices procederán de un cuerpo K arbitrario, pero fijo. (Véase el Apéndice.) Los elementos de K reciben el nombre de escalares. No se perderá nada esencial si el lector supone que K es el cuerpo real, **R**, o el complejo, **C**.

Por último, advertimos que los elementos de $\mathbf{R}^n$ o $\mathbf{C}^n$ se representarán convenientemente por «vectores fila» o «vectores columna», que son casos particulares de matrices.

3.2. MATRICES

Una *matriz sobre un cuerpo K* (o simplemente una *matriz*, si K viene dado implícitamente) es una tabla ordenada de escalares a_{ij} de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

La matriz anterior se denota también por (a_{ij}), $i = 1, \dots, m$, $j = 1, \dots, n$, o simplemente por (a_{ij}). Las m n-plas horizontales

$$(a_{11}, a_{12}, \dots, a_{1n}), (a_{21}, a_{22}, \dots, a_{2n}), \dots, (a_{m1}, a_{m2}, \dots, a_{mn})$$

son las *filas* de la matriz, y las n m-plas verticales

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}$$

son sus *columnas*. Nótese que el elemento a_{ij}, llamado *entrada ij* o *componente ij*, aparece en la fila i-ésima y en la columna j-ésima. Una matriz con m filas y n columnas se denomina matriz m por n, o matriz $m \times n$; el par de números (m, n) se llama su *tamaño* o *forma*.

Las matrices se denotarán usualmente por letras mayúsculas, A, B, ..., y los elementos del cuerpo K por minúsculas, a, b, ... Dos matrices A y B son *iguales*, escrito $A = B$, si tienen la misma forma y si sus elementos correspondientes coinciden. Así la igualdad de dos matrices $m \times n$ equivale a un sistema de mn igualdades, una por cada par de componentes.

EJEMPLO 3.1

a) La siguiente es una matriz 2×3: $\begin{pmatrix} 1 & -3 & 4 \\ 0 & 5 & -2 \end{pmatrix}$.

Sus filas son $(1, -3, 4)$ y $(0, 5, -2)$; sus columnas son $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -3 \\ 5 \end{pmatrix}$ y $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

b) La aserción $\begin{pmatrix} x+y & 2z+w \\ x-y & z-w \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}$ es equivalente al siguiente sistema de ecuaciones:

$$\begin{cases} x + y = 3 \\ x - y = 1 \\ 2z + w = 5 \\ z - w = 4 \end{cases}$$

La solución del sistema es $x = 2$, $y = 1$, $z = 3$, $w = -1$.

Nota: Para referirse a una matriz con una sola fila se utiliza también la expresión *vector fila*; y para referirse a una con una columna, la expresión *vector columna*. En particular, un elemento del cuerpo K puede verse como una matriz 1×1.

3.3. SUMA DE MATRICES Y PRODUCTO POR UN ESCALAR

Sean A y B dos matrices con el mismo tamaño (esto es, con el mismo número de filas y de columnas), digamos dos matrices $m \times n$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

La *suma* de A y B, escrito $A + B$, es la matriz obtenida sumando las entradas correspondientes de ambas:

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

El *producto* de un escalar k por la matriz A, escrito $k \cdot A$ o simplemente kA, es la matriz obtenida multiplicando cada entrada de A por k:

$$kA = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}$$

Obsérvese que $A + B$ y kA son matrices $m \times n$ también. Además definimos

$$-A = -1 \cdot A \qquad \text{y} \qquad A - B = A + (-B)$$

La suma de matrices con tamaños diferentes no está definida.

EJEMPLO 3.2. Sean $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{pmatrix}$. Entonces

$$A + B = \begin{pmatrix} 1+3 & -2+0 & 3+2 \\ 4-7 & 5+1 & -6+8 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 5 \\ -3 & 6 & 2 \end{pmatrix}$$

$$3A = \begin{pmatrix} 3 \cdot 1 & 3 \cdot (-2) & 3 \cdot 3 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot (-6) \end{pmatrix} = \begin{pmatrix} 3 & -6 & 9 \\ 12 & 15 & -18 \end{pmatrix}$$

$$2A - 3B = \begin{pmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{pmatrix} + \begin{pmatrix} -9 & 0 & -6 \\ 21 & -3 & -24 \end{pmatrix} = \begin{pmatrix} -7 & -4 & 0 \\ 29 & 7 & -36 \end{pmatrix}$$

La matriz $m \times n$ cuyas entradas son todas nulas se conoce como matriz cero y se denotará por $\mathbf{0}_{m,n}$ o simplemente por $\mathbf{0}$. Por ejemplo, la matriz cero 2×3 es

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

La matriz cero es similar al escalar 0, y se utilizará el mismo símbolo para ambos. Para cualquier matriz A,

$$A + \mathbf{0} = \mathbf{0} + A = A$$

Las propiedades básicas de las matrices, bajo las operaciones de suma matricial y producto por un escalar, se enuncian a continuación.

Teorema 3.1: Sea V el conjunto de todas las matrices $m \times n$ sobre un cuerpo K. En tal caso, para matrices arbitrarias A, B, $C \in V$ y escalares cualesquiera k_1, $k_2 \in K$,

i) $(A + B) + C = A + (B + C)$
ii) $A + \mathbf{0} = A$
iii) $A + (-A) = \mathbf{0}$
iv) $A + B = B + A$
v) $k_1(A + B) = k_1A + k_1B$
vi) $(k_1 + k_2)A = k_1A + k_2A$
vii) $(k_1k_2)A = k_1(k_2A)$
viii) $1 \cdot A = A$ y $0A = \mathbf{0}$

Utilizando las propiedades vi) y vii) se puede ver también que $A + A = 2A$, $A + A + A = 3A, \ldots$

Nota: Supongamos que los vectores en $\mathbf{R}^n$ se representan por vectores fila (o vectores columna), es decir,

$$u = (a_1, a_2, \ldots, a_n) \qquad \text{y} \qquad v = (b_1, b_2, \ldots, b_n)$$

Entonces, vistos como matrices, la suma $u + v$ y el producto ku son los siguientes:

$$u + v = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n) \qquad \text{y} \qquad ku = (ka_1, ka_2, \ldots, ka_n)$$

Pero esto corresponde precisamente a la suma y el producto por un escalar, tal y como se definieron en el Capítulo 2. Dicho de otro modo, las operaciones sobre matrices anteriores pueden verse como una generalización de las operaciones correspondientes introducidas en el Capítulo 2.

3.4. PRODUCTO DE MATRICES

El producto de dos matrices A y B, escrito AB, es algo complicado. Por esta razón, comenzaremos con un caso particular.

El producto $A \cdot B$ de una matriz fila $A = (a_i)$ y una matriz columna $B = (b_i)$, con el mismo número de elementos, se define como sigue:

$$(a_1, a_2, \ldots, a_n)\begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum_{k=1}^{n} a_kb_k$$

Nótese que $A \cdot B$ es un escalar (o matriz 1×1). El producto $A \cdot B$ no está definido si A y B tienen números diferentes de elementos.

EJEMPLO 3.3

$$(8, -4, 5)\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = 8 \cdot 3 + (-4) \cdot 2 + 5 \cdot (-1) = 24 - 8 - 5 = 11$$

Haciendo uso de lo anterior, definimos ahora el producto de matrices en general.

Definición: Supongamos que $A = (a_{ij})$ y $B = (b_{ij})$ son matrices tales que el número de columnas de A coincide con el número de filas de B; es decir, A es una matriz $m \times p$ y B una

matriz $p \times n$. Entonces el *producto* AB es la matriz $m \times n$ cuya entrada ij se obtiene multiplicando la fila i-ésima A_i de A por la columna j-ésima B^j de B:

$$AB = \begin{pmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 & \dots & A_1 \cdot B^n \\ A_2 \cdot B^1 & A_2 \cdot B^2 & \dots & A_2 \cdot B^n \\ \dots & \dots & \dots & \dots \\ A_m \cdot B^1 & A_m \cdot B^2 & \dots & A_m \cdot B^n \end{pmatrix}$$

Esto es,

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ \cdot & \dots & \cdot \\ a_{i1} & \dots & a_{ip} \\ \cdot & \dots & \cdot \\ a_{m1} & \dots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ \cdot & \dots & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \dots & \cdot \\ \cdot & \dots & \cdot & \dots & \cdot \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \cdot & \dots & \cdot \\ \cdot & c_{ij} & \cdot \\ \cdot & \dots & \cdot \\ c_{m1} & \dots & c_{mn} \end{pmatrix}$$

donde $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$.

Subrayamos el hecho de que el producto AB no está definido si A es una matriz $m \times p$ y B una matriz $q \times n$ con $p \neq q$.

EJEMPLO 3.4

a) $$\begin{pmatrix} r & s \\ t & u \end{pmatrix}\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 & ra_2 + sb_2 & ra_3 + sb_3 \\ ta_1 + ub_1 & ta_2 + ub_2 & ta_3 + ub_3 \end{pmatrix}$$

b) $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 & 1 \cdot 1 + 2 \cdot 2 \\ 3 \cdot 1 + 4 \cdot 0 & 3 \cdot 1 + 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 3 & 11 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 1 \cdot 3 & 1 \cdot 2 + 1 \cdot 4 \\ 0 \cdot 1 + 2 \cdot 3 & 0 \cdot 2 + 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 8 \end{pmatrix}$$

El ejemplo anterior muestra que el producto de matrices no es conmutativo, es decir, los productos AB y BA de matrices no son necesariamente iguales.

El producto de matrices satisface, sin embargo, las siguientes propiedades:

Teorema 3.2:
i) $(AB)C = A(BC)$ (ley asociativa).
ii) $A(B + C) = AB + AC$ (ley distributiva por la izquierda).
iii) $(B + C)A = BA + CA$ (ley distributiva por la derecha).
iv) $k(AB) = (kA)B = A(kB)$, si k es un escalar.

Se supone que las sumas y productos del teorema están definidos.

Hacemos notar que $\mathbf{0}A = \mathbf{0}$ y $B\mathbf{0} = \mathbf{0}$, donde $\mathbf{0}$ es la matriz cero.

3.5. TRASPUESTA DE UNA MATRIZ

La *traspuesta* de una matriz A, denotada por A^T, es la matriz obtenida escribiendo las filas de A, por orden, como columnas:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

En otras palabras, si $A = (a_{ij})$ es una matriz $m \times n$, entonces $A^T = (a_{ij}^T)$ es la matriz $n \times m$ con $a_{ij}^T = a_{ji}$ para todos los valores de i y j.

Adviértase que la traspuesta de un vector fila es un vector columna y viceversa.

La operación de trasposición definida sobre matrices satisface las siguientes propiedades:

Teorema 3.3: i) $(A + B)^T = A^T + B^T$.

ii) $(A^T)^T = A$.

iii) $(kA)^T = kA^T$ (si k es un escalar).

iv) $(AB)^T = B^T A^T$.

Observamos en iv) que la traspuesta de un producto es el producto de las traspuestas, pero en orden inverso.

3.6. MATRICES Y SISTEMAS DE ECUACIONES LINEALES

Consideremos nuevamente un sistema de m ecuaciones lineales con n incógnitas:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad [3.1]$$

El sistema es equivalente a la ecuación matricial:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} \quad \text{o simplemente} \quad AX = B$$

donde $A = (a_{ij})$ es la llamada *matriz de los coeficientes*, $X = (x_j)$ la columna de incógnitas y $B = (b_i)$ la columna de constantes. La afirmación de que son equivalentes significa que toda solución del sistema [3.1] es solución de la ecuación matricial, y viceversa.

La *matriz ampliada* del sistema [3.1] es la matriz:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Esto es, la matriz ampliada del sistema $AX = B$ consiste en la matriz A de coeficientes, seguida por la columna B de constantes. Observemos que el sistema [3.1] está completamente determinado por su matriz ampliada.

EJEMPLO 3.5. A continuación se dan, respectivamente, un sistema de ecuaciones lineales y su ecuación matricial equivalente.

$$\begin{aligned} 2x + 3y - 4z &= 7 \\ x - 2y - 5z &= 3 \end{aligned} \qquad \text{y} \qquad \begin{pmatrix} 2 & 3 & -4 \\ 1 & -2 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

(Nótese que el tamaño de la columna de incógnitas no es igual al tamaño de la columna de constantes.)

La matriz ampliada del sistema es

$$\begin{pmatrix} 2 & 3 & -4 & 7 \\ 1 & -2 & -5 & 3 \end{pmatrix}$$

Al estudiar ecuaciones lineales, suele ser más sencillo emplear el lenguaje y la teoría de matrices, como sugieren los siguientes teoremas.

Teorema 3.4: Supongamos que $u_1, u_2, \ldots, u_n$ son soluciones de un sistema de ecuaciones lineales homogéneo $AX = 0$. Entonces toda combinación lineal de los u_i de la forma $k_1u_1 + k_2u_2 + \cdots + k_nu_n$, donde los k_i son escalares, es también una solución de $AX = 0$. En particular, todo múltiplo ku de cualquier solución u de $AX = 0$ es también solución del sistema homogéneo.

Demostración. Se nos dice que $Au_1 = 0, Au_2 = 0, \ldots, Au_n = 0$. Por consiguiente,

$$\begin{aligned} A(ku_1 + ku_2 + \cdots + ku_n) &= k_1Au_1 + k_2Au_2 + \cdots + k_nAu_n = \\ &= k_10 + k_20 + \cdots + k_n0 = 0 \end{aligned}$$

De acuerdo con ello, $k_1u_1 + \cdots + k_nu_n$ es una solución del sistema homogéneo $AX = 0$.

Teorema 3.5: La solución general de un sistema inhomogéneo $AX = B$ puede obtenerse sumando el espacio solución W del sistema homogéneo $AX = 0$ a una solución particular v_0 del sistema inhomogéneo. (Esto es, $v_0 + W$ es la solución general de $AX = B$.)

Demostración. Sea w cualquier solución de $AX = 0$. En tal caso,

$$A(v_0 + w) = A(v_0) + A(w) = B + 0 = B$$

Es decir, la suma $v_0 + w$ es solución de $AX = B$.

Por otra parte, supongamos que v es cualquier solución de $AX = B$ (que puede ser distinta de v_0). Entonces

$$A(v - v_0) = Av - Av_0 = B - B = 0$$

Esto es, la diferencia $v - v_0$ es solución del sistema homogéneo $AX = 0$. Pero

$$v = v_0 + (v - v_0)$$

De este modo, cualquier solución de $AX = B$ puede obtenerse sumando una solución de $AX = 0$ a la solución particular v_0 de $AX = B$.

Teorema 3.6: *Supongamos que el cuerpo K es infinito (por ejemplo, si K es el cuerpo real* **R** o el complejo **C**). En tal caso, el sistema $AX = B$ no tiene solución, tiene una única solución, o tiene infinitas.

Demostración. Basta probar que si $AX = B$ tiene más de una solución, necesariamente tiene infinitas. Supongamos que u y v son soluciones distintas de $AX = B$; es decir, $Au = B$ y $Av = B$. Entonces, para todo $k \in K$,

$$A[u + k(u - v)] = Au + k(Au - Av) = B + k(B - B) = B$$

En otras palabras, para todo $k \in K$, $u + k(u - v)$ es solución de $AX = B$. Dado que todas las soluciones tales son distintas (Problema 3.21), $AX = B$ tiene un número infinito de soluciones, como se pretendía.

3.7. MATRICES POR BLOQUES

Utilizando un sistema de líneas (discontinuas) horizontales y verticales podemos partir una matriz A en otras más pequeñas llamadas *bloques* (o *celdas*) de A. La matriz A se denomina entonces una *matriz por bloques*. Obviamente, una matriz dada puede dividirse en bloques de diversas maneras; por ejemplo,

$$\begin{pmatrix} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ 3 & 1 & 4 & 5 & 9 \end{pmatrix} = \left(\begin{array}{cc|cc|c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right) = \left(\begin{array}{ccc|cc} 1 & -2 & 0 & 1 & 3 \\ \hline 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right)$$

La conveniencia de la división en bloques reside en el hecho de que el resultado de las operaciones entre matrices por bloques puede obtenerse llevando a cabo el cálculo con éstos, como si realmente fueran los elementos de las matrices. Esto se ilustra a continuación.

Supongamos que A se ha partido en bloques; digamos

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}$$

Multiplicar cada bloque por un escalar k equivale a hacerlo con cada elemento de A; así

$$kA = \begin{pmatrix} kA_{11} & kA_{12} & \dots & kA_{1n} \\ kA_{21} & kA_{22} & \dots & kA_{2n} \\ \dots & \dots & \dots & \dots \\ kA_{m1} & kA_{m2} & \dots & kA_{mn} \end{pmatrix}$$

Supongamos ahora que B es una matriz dividida en el mismo número de bloques que A; digamos

$$B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2n} \\ \dots & \dots & \dots & \dots \\ B_{m1} & B_{m2} & \dots & B_{mn} \end{pmatrix}$$

Aún más, supongamos que los bloques correspondientes de A y B tienen el mismo tamaño. Sumar dichos bloques equivale a hacerlo con los elementos correspondientes de A y B, de acuerdo con lo cual,

$$A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \dots & A_{1n} + B_{1n} \\ A_{21} + B_{21} & A_{22} + B_{22} & \dots & A_{2n} + B_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} + B_{m1} & A_{m2} + B_{m2} & \dots & A_{mn} + B_{mn} \end{pmatrix}$$

El caso del producto de matrices es menos obvio, pero aún cierto. Supongamos que dos matrices U y V se dividen en bloques como sigue

$$U = \begin{pmatrix} U_{11} & U_{12} & \dots & U_{1p} \\ U_{21} & U_{22} & \dots & U_{2p} \\ \dots & \dots & \dots & \dots \\ U_{m1} & U_{m2} & \dots & U_{mp} \end{pmatrix} \quad \text{y} \quad V = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1n} \\ V_{21} & V_{22} & \dots & V_{2n} \\ \dots & \dots & \dots & \dots \\ V_{p1} & V_{22} & \dots & V_{pn} \end{pmatrix}$$

de forma que el número de columnas en cada bloque U_{ik} es igual al número de filas en cada bloque V_{kj}. En tal caso,

$$UV = \begin{pmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \dots & \dots & \dots & \dots \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{pmatrix}$$

donde

$$W_{ij} = U_{i1}V_{1j} + U_{i2}V_{2j} + \cdots + U_{ip}V_{pj}$$

La demostración de la fórmula anterior es directa, aunque detallada y larga. Se deja como problema suplementario.

PROBLEMAS RESUELTOS

SUMA DE MATRICES Y PRODUCTO POR UN ESCALAR

3.1. Calcular:

a) $A + B$ para $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & -5 \end{pmatrix}$.

b) $3A$ y $-5A$, donde $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix}$.

a) Sumamos los elementos correspondientes:

$$A + B = \begin{pmatrix} 1+1 & 2+(-1) & 3+2 \\ 4+0 & 5+3 & 6+(-5) \end{pmatrix} = \begin{pmatrix} 2 & 1 & 5 \\ 4 & 8 & 1 \end{pmatrix}$$

b) Multiplicamos cada entrada por el escalar dado:

$$3A = \begin{pmatrix} 3\cdot 1 & 3\cdot(-2) & 3\cdot 3 \\ 3\cdot 4 & 3\cdot 5 & 3\cdot(-6) \end{pmatrix} = \begin{pmatrix} 3 & -6 & 9 \\ 12 & 15 & -18 \end{pmatrix}$$

$$-5A = \begin{pmatrix} -5\cdot 1 & -5\cdot(-2) & -5\cdot 3 \\ -5\cdot 4 & -5\cdot 5 & -5\cdot(-6) \end{pmatrix} = \begin{pmatrix} -5 & 10 & -15 \\ -20 & -25 & 30 \end{pmatrix}$$

3.2. Hallar $2A - 3B$, donde $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{pmatrix}$.

Primero efectuamos los productos por los escalares y luego la suma de matrices:

$$2A - 3B = \begin{pmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{pmatrix} + \begin{pmatrix} -9 & 0 & -6 \\ 21 & -3 & -24 \end{pmatrix} = \begin{pmatrix} -7 & -4 & 0 \\ 29 & 7 & -36 \end{pmatrix}$$

(Adviértase que multiplicamos B por -3 y después sumamos, en lugar de multiplicar B por 3 y restar. Esto suele evitar errores.)

3.3. Hallar x, y, z y w si $3\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & 6 \\ -1 & 2w \end{pmatrix} + \begin{pmatrix} 4 & x+y \\ z+w & 3 \end{pmatrix}$.

Comenzamos escribiendo cada miembro como una sola matriz:

$$\begin{pmatrix} 3x & 3y \\ 3z & 3w \end{pmatrix} = \begin{pmatrix} x+4 & x+y+6 \\ z+w-1 & 2w+3 \end{pmatrix}$$

Igualamos las entradas correspondientes entre sí obteniendo el sistema de cuatro ecuaciones

$$\begin{aligned} 3x &= x+4 \\ 3y &= x+y+6 \\ 3z &= z+w-1 \\ 3w &= 2w+3 \end{aligned} \qquad \text{o} \qquad \begin{aligned} 2x &= 4 \\ 2y &= 6+x \\ 2z &= w-1 \\ w &= 3 \end{aligned}$$

La solución es: $x = 2$, $y = 4$, $z = 1$, $w = 3$.

3.4. Demostrar el Teorema 3.1 v): Sean A y B dos matrices $m \times n$ y k un escalar. Entonces $k(A + B) = kA + kB$.

Supongamos $A = (a_{ij})$ y $B = (b_{ij})$. En tal caso, $a_{ij} + b_{ij}$ es la entrada ij de $A + B$, y por tanto $k(a_{ij} + b_{ij})$ es la de $k(A + B)$. Por otra parte, ka_{ij} y kb_{ij} son las entradas ij de kA y kB, respectivamente, y en consecuencia $ka_{ij} + kb_{ij}$ es la de $kA + kB$. Pero k, a_{ij} y b_{ij} son escalares en un cuerpo, de donde

$$k(a_{ij} + b_{ij}) = ka_{ij} + kb_{ij} \qquad \text{para todos los } i, j$$

Así $k(A + B) = kA + kB$, ya que sus entradas correspondientes son iguales.

PRODUCTO DE MATRICES

3.5. Calcular: *a*) $(3, 8, -2, 4)\begin{pmatrix} 5 \\ -1 \\ 6 \end{pmatrix}$, *b*) $(1, 8, 3, 4)(6, 1, -3, 5)$.

a) El producto no está definido cuando las matrices fila y columna tienen números diferentes de elementos.

b) El producto de dos matrices fila no está definido.

3.6. Sean $A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix}$. Hallar *a*) AB, *b*) BA.

a) Dado que A es 2×2 y B es 2×3, el producto AB está definido y es una matriz 2×3. Para obtener las entradas en la primera fila de AB, multiplicamos la primera fila $(1, 3)$ de A por las columnas $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$ y $\begin{pmatrix} -4 \\ 6 \end{pmatrix}$ de B, respectivamente:

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 3 & 1 \cdot 0 + 3 \cdot (-2) & 1 \cdot (-4) + 3 \cdot 6 \\ & & \end{pmatrix} =$$

$$= \begin{pmatrix} 2 + 9 & 0 - 6 & -4 + 18 \\ & & \end{pmatrix} = \begin{pmatrix} 11 & -6 & 14 \\ & & \end{pmatrix}$$

Para obtener las entradas en la segunda fila de AB, multiplicamos la segunda fila $(2, -1)$ de A por las columnas de B, respectivamente:

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 11 & -6 & 14 \\ 4 - 3 & 0 + 2 & -8 - 6 \end{pmatrix}$$

Así

$$AB = \begin{pmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{pmatrix}$$

b) Nótese que B es 2×3 y A es 2×2. Como los números interiores 3 y 2 no son iguales, el producto BA no está definido.

3.7. Dadas $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$, hallar *a*) AB, *b*) BA.

a) Como A es 3×2 y B es 2×3, el producto AB está definido y es una matriz 3×3. Para obtener la primera fila de AB, multiplicamos la primera de A por cada columna de B:

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 2-3 & -4-4 & -10+0 \\ & & \\ & & \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ & & \\ & & \end{pmatrix}$$

Para obtener la segunda fila de AB, multiplicamos la segunda de A por cada columna de B:

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1+0 & -2+0 & -5+0 \\ & & \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ & & \end{pmatrix}$$

Para obtener la tercera fila de AB, multiplicamos la tercera de A por cada columna de B:

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ -3+12 & 6+16 & 15+0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ 9 & 22 & 15 \end{pmatrix}$$

Así

$$AB = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ 9 & 22 & 15 \end{pmatrix}$$

b) Dado que B es 2×3 y A es 3×2, el producto BA está definido y es una matriz 2×2. Para obtener su primera fila, multiplicamos la primera de B por cada columna de A:

$$\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 2-2+15 & -1+0-20 \\ & \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ & \end{pmatrix}$$

Para obtener su segunda fila, multiplicamos la segunda de B por cada columna de A:

$$\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ 6+4+0 & -3+0+0 \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ 10 & -3 \end{pmatrix}$$

Así

$$BA = \begin{pmatrix} 15 & -21 \\ 10 & -3 \end{pmatrix}$$

Nota: Obsérvese que en este problema tanto AB como BA están definidos, pero no son iguales; de hecho, ni siquiera tienen la misma forma.

3.8. Determinar AB, siendo

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 4 & -2 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -1 & 0 & 6 \\ 1 & 3 & -5 & 1 \\ 4 & 1 & -2 & 2 \end{pmatrix}$$

Como A es 2×3 y B es 3×4, el producto está definido y es una matriz 2×4. Multiplicamos las filas de A por las columnas de B para llegar a:

$$AB = \begin{pmatrix} 4+3-4 & -2+9-1 & 0-15+2 & 12+3-2 \\ 8-2+20 & -4-6+5 & 0+10-10 & 24-2+10 \end{pmatrix} = \begin{pmatrix} 3 & 6 & -13 & 13 \\ 26 & -5 & 0 & 32 \end{pmatrix}$$

3.9. Refirámonos al Problema 3.8. Suponiendo que sólo la tercera columna del producto AB fuera de interés, ¿cómo podría calcularse de forma independiente?

Según la regla del producto de matrices, la columna j-ésima de un producto es igual al primer factor por el j-ésimo vector columna del segundo. Siendo así,

$$\begin{pmatrix} 2 & 3 & -1 \\ 4 & -2 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ -5 \\ -2 \end{pmatrix} = \begin{pmatrix} 0-15+2 \\ 0+10-10 \end{pmatrix} = \begin{pmatrix} -13 \\ 0 \end{pmatrix}$$

De forma similar, la fila i-ésima de un producto es igual al i-ésimo vector fila del primer factor por el segundo factor.

3.10. Sea A una matriz $m \times n$, con $m > 1$ y $n > 1$. Suponiendo que u y v sean vectores, discutir las condiciones bajo las cuales *a*) Au, *b*) vA está definido.

a) El producto Au está definido únicamente cuando u es un vector columna con n componentes, es decir, una matriz $n \times 1$. En tal caso, Au es un vector columna con m componentes.

b) El producto vA está definido sólo cuando v es un vector fila con m componentes, en cuyo caso vA es otro vector fila, con n componentes.

3.11. Calcular: *a*) $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}(6 \quad -4 \quad 5)$ y *b*) $(6 \quad -4 \quad 5)\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$.

a) El primer factor es 3×1 y el segundo 1×3, por lo que el producto está definido como una matriz 3×3:

$$\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}(6 \quad -4 \quad 5) = \begin{pmatrix} (2)(6) & (2)(-4) & (2)(5) \\ (3)(6) & (3)(-4) & (3)(5) \\ (-1)(6) & (-1)(-4) & (-1)(5) \end{pmatrix} = \begin{pmatrix} 12 & -8 & 10 \\ 18 & -12 & 15 \\ -6 & 4 & -5 \end{pmatrix}$$

b) El primer factor es 1×3 y el segundo 3×1, de modo que el producto está definido como una matriz 1×1, que escribimos frecuentemente como un escalar.

$$(6 \quad -4 \quad 5)\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = (12-12-5) = (-5) = -5$$

3.12. Demostrar el Teorema 3.2 i): $(AB)C = A(BC)$.

Sean $A = (a_{ij})$, $B = (b_{jk})$ y $C = (c_{kl})$. Además, sean $AB = S = (s_{ik})$ y $BC = T = (t_{jl})$. Entonces

$$s_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{im}b_{mk} = \sum_{j=1}^{m} a_{ij}b_{jk}$$

$$t_{jl} = b_{j1}c_{1l} + b_{j2}c_{2l} + \cdots + b_{jn}c_{nl} = \sum_{k=1}^{n} b_{jk}c_{kl}$$

Ahora, multiplicando S por C, esto es, AB por C, el elemento en la fila i-ésima y en la columna l-ésima de la matriz $(AB)C$ es

$$s_{i1}c_{1l} + s_{i2}c_{2l} + \cdots + s_{in}c_{nl} = \sum_{k=1}^{n} s_{ik}c_{kl} = \sum_{k=1}^{n}\sum_{j=1}^{m}(a_{ij}b_{jk})c_{kl}$$

Por otra parte, multiplicando A por T, es decir, A por BC, el elemento en la fila i-ésima y en la columna l-ésima de la matriz $A(BC)$ es

$$a_{i1}t_{1l} + a_{i2}t_{2l} + \cdots + a_{im}t_{ml} = \sum_{j=1}^{m} a_{ij}t_{jl} = \sum_{j=1}^{m}\sum_{k=1}^{n} a_{ij}(b_{jk}c_{kl})$$

Siendo iguales las sumas anteriores, el teorema queda probado.

3.13. Demostrar el Teorema 3.2 ii): $A(B+C) = AB + AC$.

Sean $A = (a_{ij})$, $B = (b_{jk})$ y $C = (c_{jk})$. Además, sean $D = B + C = (d_{jk})$, $E = AB = (e_{ik})$ y $F = AC = (f_{ik})$. Entonces

$$d_{jk} = b_{jk} + c_{jk}$$

$$e_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{im}b_{mk} = \sum_{j=1}^{m} a_{ij}b_{jk}$$

$$f_{ik} = a_{i1}c_{1k} + a_{i2}c_{2k} + \cdots + a_{im}c_{mk} = \sum_{j=1}^{m} a_{ij}c_{jk}$$

Por tanto, el elemento en la fila i-ésima y en la columna k-ésima de la matriz $AB + AC$ es

$$e_{ik} + f_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk} + \sum_{j=1}^{m} a_{ij}c_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

Por otro lado, el elemento en la fila i-ésima y en la columna k-ésima de la matriz $AD = A(B+C)$ es

$$a_{i1}d_{1k} + a_{i2}d_{2k} + \cdots + a_{im}d_{mk} = \sum_{j=1}^{m} a_{ij}d_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

Así $A(B+C) = AB + AC$ debido a que los elementos correspondientes son iguales.

TRASPUESTA

3.14. Dada $A = \begin{pmatrix} 1 & 3 & 5 \\ 6 & -7 & -8 \end{pmatrix}$, hallar A^T y $(A^T)^T$.

Reescribimos las filas de A como columnas para obtener A^T y a continuación reescribimos las filas de A^T como columnas para obtener $(A^T)^T$:

$$A^T = \begin{pmatrix} 1 & 6 \\ 3 & -7 \\ 5 & -8 \end{pmatrix} \qquad (A^T)^T = \begin{pmatrix} 1 & 3 & 5 \\ 6 & -7 & -8 \end{pmatrix}$$

[Tal y como cabía esperar por el Teorema 3.3 ii), $(A^T)^T = A$.]

3.15. Probar que las matrices AA^T y A^TA están definidas para cualquier matriz A.

Si A es una matriz $m \times n$, A^T será una matriz $n \times m$. Por consiguiente, AA^T estará definida como una matriz $m \times m$ y A^TA lo estará como una matriz $n \times n$.

3.16. Encontrar AA^T y A^TA, donde $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$.

Obtenemos A^T reescribiendo las filas de A como columnas:

$$A^T = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix} \quad \text{de donde} \quad AA^T = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 26 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 1+9 & 2-3 & 0+12 \\ 2-3 & 4+1 & 0-4 \\ 0+12 & 0-4 & 0+16 \end{pmatrix} = \begin{pmatrix} 10 & -1 & 12 \\ -1 & 5 & -4 \\ 12 & -4 & 16 \end{pmatrix}$$

3.17. Demostrar el Teorema 3.3. iv): $(AB)^T = B^TA^T$.

Si $A = (a_{ij})$ y $B = (b_{kj})$, la entrada ij de AB es

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \qquad [1]$$

De este modo, [1] es la entrada ji (orden inverso) de $(AB)^T$.

Por otra parte, la columna j de B se convierte en la fila j de B^T y la fila i de A en la columna i de A^T. Consecuentemente, la entrada ji de B^TA^T es

$$(b_{1j}, b_{2j}, \ldots, b_{mj})\begin{pmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{im} \end{pmatrix} = b_{1j}a_{i1} + b_{2j}a_{i2} + \cdots + b_{mj}a_{im}$$

Así $(AB)^T = B^TA^T$, puesto que las entradas correspondientes son iguales.

MATRICES POR BLOQUES

3.18. Calcular AB empleando multiplicación por bloques, donde

$$A = \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \quad \text{y} \quad B = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Aquí $A = \begin{pmatrix} E & F \\ 0_{1\times 2} & G \end{pmatrix}$ y $B = \begin{pmatrix} R & S \\ 0_{1\times 3} & T \end{pmatrix}$, donde E, F, G, R, S y T son los bloques dados. Por tanto,

$$AB = \begin{pmatrix} ER & ES+FT \\ 0_{1\times 3} & GT \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix} & \begin{pmatrix} 3 \\ 7 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ (0 \quad 0 \quad 0) & (2) \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 & 4 \\ 19 & 26 & 33 & 7 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

3.19. Evaluar CD mediante multiplicación por bloques, siendo

$$C = \left(\begin{array}{cc|ccc} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ \hline 0 & 0 & 5 & 1 & 2 \\ 0 & 0 & 3 & 4 & 1 \end{array}\right) \quad \text{y} \quad D = \left(\begin{array}{cc|cc} 3 & -2 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -4 & 1 \end{array}\right)$$

$$CD = \begin{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ 2 & 4 \end{pmatrix} & 0_{2\times 2} \\ 0_{2\times 2} & \begin{pmatrix} 5 & 1 & 2 \\ 3 & 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -3 \\ -4 & 1 \end{pmatrix} \end{pmatrix} =$$

$$= \begin{pmatrix} \begin{pmatrix} 3+4 & -2+8 \\ 9+8 & -6+16 \end{pmatrix} & 0_{2\times 2} \\ 0_{2\times 2} & \begin{pmatrix} 5+2-8 & 10-3+2 \\ 3+8-4 & 6-12+1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 7 & 6 & 0 & 0 \\ 17 & 10 & 0 & 0 \\ 0 & 0 & -1 & 9 \\ 0 & 0 & 7 & -5 \end{pmatrix}$$

PROBLEMAS VARIOS

3.20. Demostrar las siguientes afirmaciones: *a*) Si A tiene una fila nula, entonces AB la tiene también. *b*) Si B tiene una columna nula, también la tiene AB.

a) Sean R_i la fila nula de A y $B^1, \ldots, B^n$ las columnas de B. Entonces la fila i-ésima de AB es

$$(R_i \cdot B^1, R_i \cdot B^2, \ldots, R_i \cdot B^n) = (0, 0, \ldots, 0)$$

b) Sean C_j la columna nula de B y $A_1, \ldots, A_m$ las filas de A. En tal caso, la columna j-ésima de AB es

$$\begin{pmatrix} A_1 \cdot C_j \\ A_2 \cdot C_j \\ \cdots \\ A_m \cdot C_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}$$

3.21. Sean u y v dos vectores distintos. Probar que, para cada escalar $k \in K$, los vectores $u + k(u - v)$ son distintos.

Es suficiente probar que si

$$u + k_1(u - v) = u + k_2(u - v) \qquad [1]$$

entonces $k_1 = k_2$. Supongamos que se verifica [1]. En tal caso,

$$k_1(u - v) = k_2(u - v) \qquad \text{o} \qquad (k_1 - k_2)(u - v) = 0$$

Como u y v son distintos, $u - v \neq 0$. Por consiguiente, $k_1 - k_2 = 0$ y $k_1 = k_2$.

PROBLEMAS SUPLEMENTARIOS

OPERACIONES MATRICIALES

Los Problemas 3.22-3.25 están referidos a las siguientes matrices:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 0 \\ -6 & 7 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & -3 & 4 \\ 2 & 6 & -5 \end{pmatrix}$$

3.22. Hallar $5A - 2B$ y $2A + 3B$.

3.23. Hallar: *a*) AB y $(AB)C$, *b*) BC y $A(BC)$. [Nótese que $(AB)C = A(BC)$.]

3.24. Hallar A^T, B^T y $A^T B^T$. [Nótese que $A^T B^T \neq (AB)^T$.]

3.25. Hallar $AA = A^2$ y AC.

3.26. Supóngase que $e_1 = (1, 0, 0)$, $e_2 = (0, 0, 1)$, $e_3 = (0, 0, 1)$ y $A = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix}$. Hallar $e_1 A$, $e_2 A$ y $e_3 A$.

3.27. Sea $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, donde 1 es la componente i-ésima. Demostrar las siguientes afirmaciones:

a) $e_i A = A_i$, la i-ésima fila de una matriz A.

b) $Be_j^T = B^j$, la j-ésima columna de B.

c) Si $e_i A = e_i B$ para todo i, entonces $A = B$.

d) Si $Ae_j^T = Be_j^T$ para todo j, entonces $A = B$.

3.28. Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Encontrar una matriz 2×3 B con entradas distintas tal que $AB = 0$.

3.29. Demostrar el Teorema 3.2: iii) $(B + C)A = BA + CA$; iv) $k(AB) = (kA)B = A(kB)$, donde k es un escalar. [Las partes i) y ii) se probaron en los Problemas 3.12 y 3.13, respectivamente.]

3.30. Demostrar el Teorema 3.3: i) $(A + B)^T = A^T + B^T$; ii) $(A^T)^T = A$; iii) $(kA)^T = kA^T$, siendo k un escalar. [La parte iv) se demostró en el Problema 3.17.]

3.31. Supóngase que $A = (A_{ik})$ y $B = (B_{kj})$ son matrices por bloques para las que AB está definida y que el número de columnas de cada bloque A_{ik} es igual al número de filas de cada bloque B_{kj}. Demostrar que

$$AB = (C_{ij}), \text{ donde } C_{ij} = \sum_k A_{ik} B_{kj}$$

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

3.22. $\begin{pmatrix} -5 & 10 \\ 27 & -36 \end{pmatrix}, \begin{pmatrix} 17 & 4 \\ -12 & 13 \end{pmatrix}$

3.23. *a*) $\begin{pmatrix} -7 & 14 \\ 39 & -28 \end{pmatrix}, \begin{pmatrix} 21 & 105 & -98 \\ -17 & -285 & 296 \end{pmatrix}$; *b*) $\begin{pmatrix} 5 & -15 & 20 \\ 8 & 60 & -59 \end{pmatrix}, \begin{pmatrix} 21 & 105 & -98 \\ -17 & -285 & 296 \end{pmatrix}$

3.24. $\begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix}, \begin{pmatrix} 5 & -6 \\ 0 & 7 \end{pmatrix}, \begin{pmatrix} 5 & 15 \\ 10 & -40 \end{pmatrix}$

3.25. $\begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix}, \begin{pmatrix} 5 & 9 & -6 \\ -5 & -33 & 32 \end{pmatrix}$

3.26. (a_1, a_2, a_3, a_4), (b_1, b_2, b_3, b_4), (c_1, c_2, c_3, c_4), las filas de A.

3.28. $\begin{pmatrix} 2 & 4 & 6 \\ -1 & -2 & -3 \end{pmatrix}$

CAPITULO 4

Matrices cuadradas. Matrices elementales

4.1. INTRODUCCION

Las matrices con el mismo número de filas que de columnas se denominan *matrices cuadradas*. Son de importancia capital dentro del álgebra lineal y se emplearán a lo largo de todo el texto. Este capítulo introduce dichas matrices junto a un cierto número de sus propiedades básicas.

Además, el capítulo nos presenta las matrices elementales, que están estrechamente ligadas a las operaciones elementales entre filas del Capítulo 1. Las utilizaremos para justificar dos algoritmos: uno que determina la inversa de una matriz, y otro que diagonaliza una forma cuadrática.

A menos que se establezca o sobrentienda lo contrario, los escalares serán números reales en este capítulo. No obstante, discutiremos el caso especial de las matrices complejas y algunas de sus propiedades.

4.2. MATRICES CUADRADAS

Una *matriz cuadrada* es la que tiene el mismo número de filas que de columnas. Se dice que una matriz cuadrada $n \times n$ es de *orden* n y se le asigna el nombre de *matriz n-cuadrada*.

Recuérdese que no todo par de matrices puede sumarse o multiplicarse. Sin embargo, cuando nos limitamos a considerar las de un orden dado, n, este inconveniente desaparece. Especificando, las operaciones de suma, producto, producto por un escalar y trasposición pueden efectuarse sobre todas las matrices $n \times n$ y el resultado es nuevamente una matriz $n \times n$.

EJEMPLO 4.1. Sean $A = \begin{pmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{pmatrix}$. Entonces A y B son matrices cuadradas de orden 3. Además,

$$A+B=\begin{pmatrix}3 & -3 & 4\\ -4 & -1 & -6\\ 6 & 8 & 3\end{pmatrix}\qquad 2A=\begin{pmatrix}2 & 4 & 6\\ -8 & -8 & -8\\ 10 & 12 & 14\end{pmatrix}\qquad A^T=\begin{pmatrix}1 & -4 & 5\\ 2 & -4 & 6\\ 3 & -4 & 7\end{pmatrix}$$

y

$$AB=\begin{pmatrix}2+0+3 & -5+6+6 & 1-4-12\\ -8+0-4 & 20-12-8 & -4+8+16\\ 10+0+7 & -25+18+14 & 5-12-28\end{pmatrix}=\begin{pmatrix}5 & 7 & -15\\ -12 & 0 & 20\\ 17 & 7 & -35\end{pmatrix}$$

son matrices de orden 3.

Nota: Una colección no vacía **A** de matrices se denomina un *álgebra* (de matrices) si es cerrada bajo las operaciones de suma, producto por un escalar y producto. Así la colección $\mathbf{M}_n$ de todas las matrices n-cuadradas constituye un álgebra de matrices.

MATRICES CUADRADAS COMO FUNCIONES

Sea A cualquier matriz n-cuadrada. A puede verse como una función $A:\mathbf{R}^n \to \mathbf{R}^n$ de dos formas diferentes:

1. $A(u) = Au$, donde u es un vector columna.
2. $A(u) = uA$, donde u es un vector fila.

En este texto se adopta el primer significado de $A(u)$, esto es, la función definida por la matriz A será $A(u) = Au$. De acuerdo con esto, a menos que se establezca o sobrentienda lo contrario, se supondrá que los vectores u en $\mathbf{R}^n$ son vectores columna (no vectores fila). Por conveniencia tipográfica, tales vectores columna u se presentarán a menudo como vectores fila traspuestos.

EJEMPLO 4.2. Sea $A=\begin{pmatrix}1 & -2 & 3\\ 4 & 5 & -6\\ 2 & 0 & -1\end{pmatrix}$. Si $u = (1, -3, 7)^T$, entonces

$$A(u)=Au=\begin{pmatrix}1 & -2 & 3\\ 4 & 5 & -6\\ 2 & 0 & -1\end{pmatrix}\begin{pmatrix}1\\ -3\\ 7\end{pmatrix}=\begin{pmatrix}1+6+21\\ 4-15-42\\ 2+0-7\end{pmatrix}=\begin{pmatrix}28\\ -53\\ -5\end{pmatrix}$$

Si $w = (2, -1, 4)^T$, entonces

$$A(w)=Aw=\begin{pmatrix}1 & -2 & 3\\ 4 & 5 & -6\\ 2 & 0 & -1\end{pmatrix}\begin{pmatrix}2\\ -1\\ 4\end{pmatrix}=\begin{pmatrix}2+2+12\\ 8-5-24\\ 4+0-4\end{pmatrix}=\begin{pmatrix}16\\ -21\\ 0\end{pmatrix}$$

Advertencia: En algunos textos se define la función A anterior como

$$A(u) = uA$$

donde se supone que los vectores u en $\mathbf{R}^n$ son vectores fila. Evidentemente, si uno sólo trabaja con vectores en $\mathbf{R}^n$ y no con matrices, carece de importancia la forma en que se definan éstos.

MATRICES QUE CONMUTAN

Se dice que las matrices A y B *conmutan* si $AB = BA$, condición sólo verificable por matrices cuadradas del mismo orden. Por ejemplo, supongamos

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 5 & 4 \\ 6 & 11 \end{pmatrix}$$

Entonces

$$AB = \begin{pmatrix} 5+12 & 4+22 \\ 15+24 & 12+44 \end{pmatrix} = \begin{pmatrix} 17 & 26 \\ 39 & 56 \end{pmatrix} \quad \text{y} \quad BA = \begin{pmatrix} 5+12 & 10+16 \\ 6+33 & 12+44 \end{pmatrix} = \begin{pmatrix} 17 & 26 \\ 39 & 56 \end{pmatrix}$$

Como $AB = BA$, las matrices conmutan.

4.3. DIAGONAL Y TRAZA. MATRIZ IDENTIDAD

Sea $A = (a_{ij})$ una matriz n-cuadrada. La *diagonal* (o *diagonal principal*) de A consiste en los elementos $a_{11}, a_{22}, \ldots, a_{nn}$. La *traza* de A, escrito tr A, es la suma de los elementos diagonales, es decir,

$$\text{tr } A = a_{11} + a_{22} + \cdots + a_{nn} \equiv \sum_{i=1}^{n} a_{ii}$$

La matriz n-cuadrada con unos en la diagonal y ceros en cualquier otra posición, denotada por I_n o simplemente por I, se conoce como *matriz identidad* (o *unidad*). La matriz I se asemeja al escalar 1 en que, para cualquier matriz A (del mismo orden),

$$AI = IA = A$$

Con mayor generalidad, si B es una matriz $m \times n$, $BI_n = B$ e $I_m B = B$ (Problema 4.9).

Para un escalar arbitrario $k \in K$, la matriz kI, que contiene k en las posiciones diagonales y 0 en cualquier otro lugar, se denomina la *matriz escalar* correspondiente al escalar k.

EJEMPLO 4.3

a) La delta de Kronecker δ_{ij} se define como

$$\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Siendo así, la matriz identidad puede definirse como $I = (\delta_{ij})$.

b) Las matrices escalares de órdenes 2, 3 y 4 correspondientes al escalar $k = 5$ son, respectivamente,

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 5 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{pmatrix}$$

(Es práctica común el omitir bloques o disposiciones de ceros como se ha hecho en la tercera matriz.)

El siguiente teorema se demostrará en el Problema 4.10.

Teorema 4.1: Supongamos que $A = (a_{ij})$ y $B = (b_{ij})$ son matrices n-cuadradas y que k es un escalar. En tal caso,

i) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$, ii) $\operatorname{tr} kA = k \cdot \operatorname{tr} A$, iii) $\operatorname{tr} AB = \operatorname{tr} BA$.

4.4. POTENCIAS DE MATRICES. POLINOMIOS DE MATRICES

Sea A una matriz n-cuadrada sobre un cuerpo K. Las potencias de A se definen como sigue:

$$A^2 = AA \qquad A^3 = A^2 A, \ldots, A^{n+1} = A^n A, \ldots \qquad \text{y} \qquad A^0 = I$$

También se definen polinomios en la matriz A. Concretamente, para cualquier polinomio,

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

donde los a_i son escalares, $f(A)$ se define como la matriz

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n$$

[Nótese que $f(A)$ se obtiene de $f(x)$ sustituyendo la variable x por la matriz A y el escalar a_0 por la matriz escalar $a_0 I$.] En caso de que $f(A)$ sea la matriz cero, A recibe el nombre de *cero* o *raíz* del polinomio $f(x)$.

EJEMPLO 4.4. Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$. Entonces

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} \qquad \text{y} \qquad A^3 = A^2 A = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} -11 & 38 \\ 57 & -106 \end{pmatrix}$$

Si $f(x) = 2x^2 - 3x + 5$, entonces

$$f(A) = 2\begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} - 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 16 & -18 \\ -27 & 61 \end{pmatrix}$$

Si $g(x) = x^2 + 3x - 10$, entonces

$$g(A) = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} + 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} - 10\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Así A es un cero del polinomio $g(x)$.

La aplicación anterior del anillo $K[x]$ de los polinomios sobre K en el álgebra $\mathbf{M}_n$ de las matrices n-cuadradas definida por

$$f(x) \to f(A)$$

se llama *aplicación de evaluación en A*.

Es aplicable el teorema enunciado a continuación (y probado en el Problema 4.11).

Teorema 4.2: Sean $f(x)$ y $g(x)$ dos polinomios y A una matriz n-cuadrada (todos sobre K). Se cumple:

i) $(f+g)(A) = f(A) + g(A)$.

ii) $(fg)(A) = f(A)g(A)$.

iii) $f(A)g(A) = g(A)f(A)$.

Dicho de otro modo, i) y ii) establecen que si primero sumamos (multiplicamos) los polinomios $f(x)$ y $g(x)$ y luego evaluamos la suma (producto) en A, el resultado coincide con el que obtendríamos si primero evaluáramos $f(x)$ y $g(x)$ en A y después sumáramos (multiplicáramos) las matrices $f(A)$ y $g(A)$. La parte iii) establece que dos polinomios cualesquiera en A conmutan.

4.5. MATRICES INVERTIBLES (NO SINGULARES)

Se dice que una matriz cuadrada A es *invertible* (o *no singular*) si existe una matriz B con la propiedad de que

$$AB = BA = I$$

siendo I la matriz identidad. Tal matriz B es única, puesto que

$$AB_1 = B_1A = I \quad \text{y} \quad AB_2 = B_2A = I \qquad \text{implica} \qquad B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$$

Denominamos a la matriz B la *inversa* de A y la denotamos por A^{-1}. Observemos que la relación anterior es simétrica, o sea, si B es la inversa de A, necesariamente A es la inversa de B.

EJEMPLO 4.5

a) Supongamos $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$. Entonces

$$AB = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$BA = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 6-5 & 15-15 \\ -2+2 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

De este modo, A y B son invertibles, siendo cada una la inversa de la otra.

b) Supongamos $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$ y $B = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$. Entonces

$$AB = \begin{pmatrix} -11+0+12 & 2+0-2 & 2+0-2 \\ -22+4+18 & 4+0-3 & 4-1-3 \\ -44-4+48 & 8+0-8 & 8+1-8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Por el Problema 4.21, $AB = I$ si y sólo si $BA = I$; por consiguiente, no necesitamos comprobar si $BA = I$. Siendo así, A y B son cada una inversa de la otra.

Consideremos ahora una matriz 2×2 general

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Podemos determinar cuándo es A invertible y, en caso de serlo, dar una fórmula para su inversa. Para empezar buscamos escalares x, y, z, t tales que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} ax + bz & ay + bt \\ cx + dz & cy + dt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

lo cual se reduce a resolver los dos sistemas:

$$\begin{cases} ax + bz = 1 \\ cx + dz = 0 \end{cases} \qquad \begin{cases} ay + bt = 0 \\ cy + dt = 1 \end{cases}$$

donde la matriz original A es la matriz de los coeficientes de cada sistema. Tomemos $|A| = ad - bc$ (el determinante de A). Según los Problemas 1.60 y 1.61, los dos sistemas son resolubles y por ende A es invertible si y sólo si $|A| \neq 0$. En ese caso, el primer sistema tiene la solución única $x = d/|A|$, $z = -c/|A|$ y el segundo la solución única $y = -b/|A|$, $t = a/|A|$. De acuerdo con esto,

$$A^{-1} = \begin{pmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Expresado en palabras: Cuando $|A| \neq 0$, la inversa de una matriz 2×2 A se obtiene: i) intercambiando los elementos de la diagonal principal, ii) tomando los opuestos de los otros elementos, iii) multiplicando la matriz resultante por $1/|A|$.

Nota 1: La propiedad de que A es invertible si y sólo si su determinante $|A| \neq 0$ es cierta para matrices cuadradas de cualquier orden. (Véase Capítulo 7.)

Nota 2: Supóngase que A y B son invertibles. En tal caso, AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$. Con mayor generalidad, si $A_1, A_2, \ldots, A_k$ son invertibles, entonces su producto también lo es y

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1} A_1^{-1}$$

que es el producto de las inversas en orden inverso.

4.6. TIPOS ESPECIALES DE MATRICES CUADRADAS

Esta sección describe un cierto número de clases especiales de matrices que desempeñan un papel importante en el álgebra lineal.

MATRICES DIAGONALES

Una matriz cuadrada $D = (d_{ij})$ es *diagonal* si todas sus entradas no diagonales son nulas. Tal matriz se denota frecuentemente por $D = \text{diag}\,(d_{11}, d_{22}, \ldots, d_{nn})$, donde algunos o todos los d_{ii} pueden ser nulos. Por ejemplo,

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 4 & 0 \\ 0 & -5 \end{pmatrix} \qquad \begin{pmatrix} 6 & & & \\ & 0 & & \\ & & -9 & \\ & & & 1 \end{pmatrix}$$

son matrices diagonales que pueden representarse, respectivamente, por

$$\text{diag}\,(3,\ -7,\ 2) \qquad \text{diag}\,(4,\ -5) \qquad \text{y} \qquad \text{diag}\,(6,\ 0,\ -9,\ 1)$$

(Obsérvese que se han omitido disposiciones de ceros en la tercera matriz.)

Claramente, la suma, producto por un escalar y producto de matrices diagonales son nuevamente diagonales. De esto modo, todas las matrices diagonales $n \times n$ forman un álgebra de matrices. De hecho, las matrices diagonales forman un álgebra conmutativa, porque dos matrices diagonales $n \times n$ cualesquiera conmutan.

MATRICES TRIANGULARES

Una matriz cuadrada $A = (a_{ij})$ es una *matriz triangular superior* o simplemente una *matriz triangular* si todas las entradas bajo la diagonal principal son iguales a cero, esto es, si $s_{ij} = 0$ para $i > j$. Las matrices triangulares superiores genéricas de órdenes 2, 3 y 4 son, respectivamente,

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \qquad \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ & b_{22} & b_{23} \\ & & b_{33} \end{pmatrix} \qquad \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ & c_{22} & c_{23} & c_{24} \\ & & c_{33} & c_{34} \\ & & & c_{44} \end{pmatrix}$$

(Como con las matrices diagonales, es práctica común el omitir disposiciones de ceros.)

Las matrices triangulares superiores también forman un álgebra de matrices. De hecho,

Teorema 4.3: Supongamos que $A = (a_{ij})$ y $B = (b_{ij})$ son matrices triangulares superiores. Entonces

i) $A + B$ es triangular superior, con $\text{diag}\,(a_{11} + b_{11}, a_{22} + b_{22}, \ldots, a_{nn} + b_{nn})$.

ii) kA es triangular superior, con $\text{diag}\,(ka_{11}, ka_{22}, \ldots, ka_{nn})$.

iii) AB es triangular superior, con $\text{diag}\,(a_{11}b_{11}, a_{22}b_{22}, \ldots, a_{nn}b_{nn})$.

iv) Para cualquier polinomio $f(x)$, la matriz $f(A)$ es triangular superior, con

$$\text{diag}\,(f(a_{11}),\ f(a_{22}),\ \ldots,\ f(a_{nn}))$$

v) A es invertible si y sólo si cada elemento diagonal $a_{ii} \neq 0$.

Análogamente, una *matriz triangular inferior* es una matriz cuadrada cuyas entradas sobre la diagonal principal son todas nulas, y un teorema análogo al 4.3 rige para tales matrices.

MATRICES SIMETRICAS

Una matriz real es *simétrica* si $A^T = A$. Equivalentemente, $A = (a_{ij})$ es simétrica si los *elementos simétricos* (imágenes especulares respecto a la diagonal) son iguales, es decir, si cada $a_{ij} = a_{ji}$. (Nótese que A debe ser cuadrada para que pueda ser $A^T = A$.)

Una matriz real A es *antisimétrica* si $A^T = -A$. Equivalentemente, una matriz $A = (a_{ij})$ es antisimétrica si cada $a_{ij} = -a_{ji}$. Claramente, los elementos diagonales de una matriz antisimétrica deben ser nulos, ya que $a_{ii} = -a_{ii}$ implica $a_{ii} = 0$.

EJEMPLO 4.6. Consideremos las siguientes matrices:

$$A = \begin{pmatrix} 2 & -3 & 5 \\ -3 & 6 & 7 \\ 5 & 7 & -8 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 3 & -4 \\ -3 & 0 & 5 \\ 4 & -5 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

a) Por simple inspección, vemos que los elementos simétricos de A son iguales, o que $A^T = A$. Siendo así, A es simétrica.

b) Por inspección, vemos que los elementos diagonales de B son 0 y que los elementos simétricos son opuestos entre sí. De este modo, B es antisimétrica.

c) Como C no es cuadrada, no es simétrica ni antisimétrica.

Si A y B son matrices simétricas, $A + B$ y kA también lo son. Sin embargo, AB no es necesariamente simétrica. Por ejemplo,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \text{ y } B = \begin{pmatrix} 4 & 5 \\ 5 & 6 \end{pmatrix} \text{ son simétricas, pero } AB = \begin{pmatrix} 14 & 17 \\ 23 & 28 \end{pmatrix} \text{ no es simétrica}$$

Así las matrices simétricas no forman un álgebra de matrices.

El siguiente teorema se demostrará en el Problema 4.29.

Teorema 4.4: Si A es una matriz cuadrada, entonces i) $A + A^T$ es simétrica; ii) $A - A^T$ es antisimétrica; iii) $A = B + C$ para alguna matriz simétrica B y alguna antisimétrica C.

MATRICES ORTOGONALES

Se dice que una matriz real A es *ortogonal* si $AA^T = A^T A = I$. Observemos que una matriz ortogonal A es necesariamente cuadrada e invertible, con inversa $A^{-1} = A^T$.

EJEMPLO 4.7. Sea $A = \begin{pmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}$. Entonces

$$AA^T = \begin{pmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix} \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & \frac{8}{9} \\ \frac{8}{9} & -\frac{4}{9} & \frac{1}{9} \\ -\frac{4}{9} & -\frac{7}{9} & \frac{4}{9} \end{pmatrix} = \frac{1}{81} \begin{pmatrix} 1+64+16 & 4-32+28 & 8+8-16 \\ 4-32+28 & 16+16+49 & 32-4-28 \\ 8+\ 8-16 & 32-\ 4-28 & 64+1+16 \end{pmatrix} =$$

$$= \frac{1}{81} \begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Esto significa que $A^T = A^{-1}$ y por tanto $A^T A = I$. Así A es ortogonal.

Consideremos ahora una matriz 3×3 arbitraria

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Si A es ortogonal,

$$AA^T = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Esto proporciona

$$\begin{array}{rrr} a_1^2 + a_2^2 + a_3^2 = 1 & a_1b_1 + a_2b_2 + a_3b_3 = 0 & a_1c_1 + a_2c_2 + a_3c_3 = 0 \\ b_1a_1 + b_2a_2 + b_3a_3 = 0 & b_1^2 + b_2^2 + b_3^2 = 1 & b_1c_1 + b_2c_2 + b_3c_3 = 0 \\ c_1a_1 + c_2a_2 + c_3a_3 = 0 & c_1b_1 + c_2b_2 + c_3b_3 = 0 & c_1^2 + c_2^2 + c_3^2 = 1 \end{array}$$

o, en otras palabras,

$$\begin{array}{lll} u_1 \cdot u_1 = 1 & u_1 \cdot u_2 = 0 & u_1 \cdot u_3 = 0 \\ u_2 \cdot u_1 = 0 & u_2 \cdot u_2 = 1 & u_2 \cdot u_3 = 0 \\ u_3 \cdot u_1 = 0 & u_3 \cdot u_2 = 0 & u_3 \cdot u_3 = 1 \end{array}$$

donde $u_1 = (a_1, a_2, a_3)$, $u_2 = (b_1, b_2, b_3)$, $u_3 = (c_1, c_2, c_3)$ son las filas de A. Así las filas u_1, u_2 y u_3 son ortogonales entre sí y tienen longitudes unidad, o, dicho de otro modo, forman un conjunto ortonormal de vectores. La condición $A^T A = I$ muestra, similarmente, que las columnas de A constituyen un conjunto ortonormal de vectores. Aún más, dado que cada paso es reversible, el recíproco es cierto.

El resultado relativo a matrices 3×3 anterior es válido en general. Esto es,

Teorema 4.5: Sea A una matriz real. Son equivalentes las afirmaciones: *a*) A es ortogonal; *b*) las filas de A forman un conjunto ortonormal; *c*) las columnas de A forman un conjunto ortonormal.

Para $n = 2$ tenemos el resultado siguiente, probado en el Problema 4.33.

Teorema 4.6: Toda matriz ortogonal 2 × 2 tiene la forma $\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$ o $\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$ para algún número real θ.

Nota: La condición de que los vectores $u_1, u_2, \ldots, u_m$ formen un conjunto ortonormal puede describirse sencillamente como $u_i \cdot u_j = \delta_{ij}$, donde δ_{ij} es la delta de Kronecker [Ejemplo 4.3 *a*)].

MATRICES NORMALES

Una matriz real A es *normal* si conmuta con su traspuesta, esto es, si $AA^T = A^T A$. Obviamente, si A es simétrica, ortogonal o antisimétrica, es necesariamente normal. No obstante, éstas no son las únicas matrices normales.

EJEMPLO 4.8. Sea $A = \begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix}$. Entonces

$$AA^T = \begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} 6 & 3 \\ -3 & 6 \end{pmatrix} = \begin{pmatrix} 45 & 0 \\ 0 & 45 \end{pmatrix} \quad \text{y} \quad A^T A = \begin{pmatrix} 6 & 3 \\ -3 & 6 \end{pmatrix}\begin{pmatrix} 6 & -3 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 45 & 0 \\ 0 & 45 \end{pmatrix}$$

Como $AA^T = A^T A$, la matriz A es normal.

El teorema que sigue, demostrado en el Problema 4.35, caracteriza completamente las matrices reales normales 2 × 2.

Teorema 4.7: Sea A una matriz real normal 2 × 2. En tal caso, A es bien simétrica o bien la suma de una matriz escalar y otra antisimétrica.

4.7. MATRICES COMPLEJAS

Sea A una matriz compleja, es decir, una matriz cuyas entradas son números complejos. Recordemos (Sección 2.9) que si $z = a + bi$ es un número complejo, $\bar{z} = a - bi$ es su conjugado. La conjugada de la matriz compleja A, escrito $\bar{A}$, es la que se obtiene de A tomando el conjugado de cada una de sus entradas, esto es, si $A = (a_{ij})$, entonces $\bar{A} = (b_{ij})$, donde $b_{ij} = \overline{a_{ij}}$. [Denotamos este hecho escribiendo $\bar{A} = (\overline{a_{ij}})$.]

Las operaciones de trasposición y conjugación conmutan para cualquier matriz compleja A, es decir, $(\bar{A})^T = (\overline{A^T})$. De hecho, para la traspuesta conjugada de A, se utiliza la notación especial A^H. (Nótese que si A es real, $A^H = A^T$.)

EJEMPLO 4.9. Sea $A = \begin{pmatrix} 2+8i & 5-3i & 4-7i \\ 6i & 1-4i & 3+2i \end{pmatrix}$. Entonces

$$A^H = \begin{pmatrix} \overline{2+8i} & \overline{6i} \\ \overline{5-3i} & \overline{1-4i} \\ \overline{4-7i} & \overline{3+2i} \end{pmatrix} = \begin{pmatrix} 2-8i & -6i \\ 5+3i & 1+4i \\ 4+7i & 3-2i \end{pmatrix}$$

MATRICES COMPLEJAS HERMITICAS, UNITARIAS Y NORMALES

Se dice que una matriz cuadrada compleja A es *hermítica* o *antihermítica* según sea

$$A^H = A \qquad \text{o} \qquad A^H = -A$$

Si $A = (a_{ij})$ es hermítica, $a_{ij} = \overline{a_{ji}}$ y por tanto cada elemento diagonal a_{ii} debe ser real. De forma similar, si A es antihermítica, cada elemento diagonal $a_{ii} = 0$.

Se dice que una matriz cuadrada compleja A es *unitaria* si

$$A^H = A^{-1}$$

Una matriz compleja A es unitaria si y sólo si sus filas (columnas) forman un conjunto ortonormal de vectores respecto al producto interno de vectores complejos. (Véase el Problema 4.39.)

Nótese que cuando A es real, hermítica es lo mismo que simétrica y unitaria es lo mismo que ortogonal.

Se dice que una matriz cuadrada compleja A es *normal* si

$$AA^H = A^H A$$

Esta definición se reduce a la dada para matrices reales cuando A es real.

EJEMPLO 4.10: Consideremos las siguientes matrices:

$$A = \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1-2i & 4+7i \\ 1+2i & -4 & -2i \\ 4-7i & 2i & 2 \end{pmatrix} \qquad C = \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix}$$

a) A es unitaria si $A^H = A^{-1}$ o si $AA^H = A^H A = I$. Tal y como se indicó previamente, sólo es necesario probar que $AA^H = I$:

$$AA^H = A\bar{A}^T = \frac{1}{4}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}\begin{pmatrix} 1 & -i & 1-i \\ i & 1 & -1-i \\ -1-i & 1-i & 0 \end{pmatrix} =$$

$$= \frac{1}{4}\begin{pmatrix} 1+1+2 & -i-i+2i & 1-i+i-1+0 \\ i+i-2i & 1+1+2 & i+1-1-i \\ 1+i-i-1+0 & -i+1-1+i+0 & 2+2+0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

De acuerdo con esto, A es unitaria.

b) B es hermítica, puesto que sus elementos diagonales 3, -4 y 2 son reales, y los elementos simétricos $1-2i$ y $1+2i$, $4+7i$ y $4-7i$, y $-2i$ y $2i$ son conjugados.

c) Para probar que C es normal, evaluamos CC^H y $C^H C$:

$$CC^H = C\bar{C}^T = \begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix}\begin{pmatrix} 2-3i & -i \\ 1 & 1-2i \end{pmatrix} = \begin{pmatrix} 14 & 4-4i \\ 4+4i & 6 \end{pmatrix}$$

$$C^H C = \bar{C}^T C = \begin{pmatrix} 2-3i & -i \\ 1 & 1-2i \end{pmatrix}\begin{pmatrix} 2+3i & 1 \\ i & 1+2i \end{pmatrix} = \begin{pmatrix} 14 & 4-4i \\ 4+4i & 6 \end{pmatrix}$$

Como $CC^H = C^H C$, la matriz compleja C es normal.

4.8. MATRICES CUADRADAS POR BLOQUES

Una matriz por bloques A se llama *matriz cuadrada por bloques* si: i) A es cuadrada, ii) los bloques forman una matriz cuadrada, iii) los bloques diagonales son también matrices cuadradas. Las dos últimas condiciones se darán si y sólo si hay el mismo número de líneas horizontales y verticales y están situadas simétricamente.

Consideremos las dos matrices por bloques:

$$A = \left(\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 4 & 4 & 4 & 4 & 4 \\ 3 & 5 & 3 & 5 & 3 \end{array}\right) \qquad B = \left(\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ 4 & 4 & 4 & 4 & 4 \\ \hline 3 & 5 & 3 & 5 & 3 \end{array}\right)$$

A no es una matriz cuadrada por bloques debido a que los bloques diagonales segundo y tercero no son matrices cuadradas. Por el contrario, B es una matriz cuadrada por bloques.

Una *matriz diagonal por bloques* M es una matriz cuadrada por bloques en la que todos los no diagonales son matrices cero. La importancia de estas matrices radica en el hecho de que el álgebra de las matrices por bloques se reduce frecuentemente a la de los bloques individuales. Específicamente, supongamos que M es una matriz diagonal por bloques y que $f(x)$ es cualquier polinomio. En tal caso, M y $f(M)$ tienen la siguiente forma:

$$M = \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \cdots & \\ & & & A_{rr} \end{pmatrix} \qquad f(M) = \begin{pmatrix} f(A_{11}) & & & \\ & f(A_{22}) & & \\ & & \cdots & \\ & & & f(A_{rr}) \end{pmatrix}$$

(Como de costumbre, empleamos espacios en blanco para disposiciones o bloques de ceros.)

Análogamente, una matriz cuadrada por bloques recibe el nombre de *matriz triangular superior por bloques* si los bloques bajo la diagonal son matrices cero, y el de *matriz triangular inferior por bloques* si lo son los bloques sobre la diagonal.

4.9. MATRICES ELEMENTALES. APLICACIONES

Comencemos recordando (Sección 1.8) las llamadas *operaciones elementales entre filas* en una matriz A:

$[E_1]$ (Intercambio de filas) Intercambiar las filas i-ésima y j-ésima:

$$R_i \leftrightarrow R_j$$

$[E_2]$ (Cambio de escala de filas) Multiplicar la fila i-ésima por un escalar no nulo k:

$$kR_i \to R_i \qquad k \neq 0$$

$[E_3]$ (Adición de filas) Sustituir la fila i-ésima por ella misma más k veces la j-ésima:

$$kR_j + R_i \to R_i$$

Cada una de las operaciones anteriores tiene una inversa del mismo tipo. De forma específica (Problema 4.19):

1. $R_j \rightarrow R_i$ es su propia inversa.
2. $kR_i \rightarrow R_i$ y $k^{-1}R_i \rightarrow R_i$ son inversas.
3. $kR_j + R_i \rightarrow R_i$ y $-kR_j + R_i \rightarrow R_i$ son inversas.

Recordemos también (Sección 1.8) que se dice que una matriz B es *equivalente por filas* a otra A, escrito $A \sim B$, si B puede obtenerse de A mediante una sucesión finita de operaciones elementales entre filas. Dado que éstas son invertibles, la equivalencia por filas es una relación de equivalencia; es decir: *a*) $A \sim A$; *b*) si $A \sim B$, entonces $B \sim A$; *c*) si $A \sim B$ y $B \sim C$, necesariamente $A \sim C$. Asimismo, establecemos de nuevo el siguiente resultado básico, relativo a la equivalencia por filas:

Teorema 4.8: Toda matriz A es equivalente por filas a una única matriz en forma canónica por filas.

MATRICES ELEMENTALES

Denotemos por e una operación elemental entre filas y por $e(A)$ el resultado de efectuarla sobre una matriz A. La matriz E obtenida efectuando e sobre la matriz identidad,

$$E = e(I)$$

se llama la *matriz elemental* correspondiente a la operación elemental entre filas e.

EJEMPLO 4.11. Las matrices 3-cuadradas elementales correspondientes a las operaciones elementales entre filas $R_2 \leftrightarrow R_3$, $-6R_2 \rightarrow R_2$ y $-4R_1 + R_3 \rightarrow R_3$ son, respectivamente,

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

El siguiente teorema, probado en el Problema 4.18, muestra la relación fundamental entre las operaciones elementales entre filas y sus matrices elementales correspondientes.

Teorema 4.9: Sean e una operación elemental entre filas y E la matriz m-cuadrada elemental correspondiente, es decir, $E = e(I_m)$. En tal caso, para cualquier matriz $m \times nA$, $e(A) = EA$.

Esto es, el resultado de efectuar una operación elemental entre filas e sobre una matriz A puede obtenerse multiplicando por la izquierda A por la matriz elemental correspondiente E.

Ahora supongamos que e' es la inversa de una operación elemental entre filas e. Sean E y E' las matrices correspondientes. Probaremos en el Problema 4.19 que E es invertible y que E' es su inversa. Esto quiere decir, en particular, que cualquier producto

$$P = E_k \cdots E_2 E_1$$

de matrices elementales es no singular.

Haciendo uso del Teorema 4.9 podemos, asimismo, demostrar (Problema 4.20) el resultado fundamental, referente a matrices invertibles, que sigue:

Teorema 4.10: Sea A una matriz cuadrada. Entonces son equivalentes las aserciones:

i) A es invertible (no singular).
ii) A es equivalente por filas a la matriz identidad I.
iii) A es producto de matrices elementales.

También utilizaremos el Teorema 4.9 para probar los que se enuncian a continuación:

Teorema 4.11: Si $AB = I$, necesariamente $BA = I$ y por consiguiente $B = A^{-1}$.

Teorema 4.12: B es equivalente por filas a A si y sólo si existe una matriz no singular P tal que $B = PA$.

APLICACION AL CALCULO DE INVERSAS

Supongamos que una matriz A es invertible y, digamos, es reducible por filas a la matriz identidad I mediante la sucesión de operaciones elementales $e_1, e_2, \ldots, e_q$. Sea E_i la matriz elemental correspondiente a la operación e_i. Según el Teorema 4.9,

$$E_q \cdots E_2 E_1 A = I \qquad \text{y} \qquad (E_q \cdots E_2 E_1 I)A = I \qquad \text{por lo que} \qquad A^{-1} = E_q \cdots E_2 E_1 I$$

En otras palabras, A^{-1} puede obtenerse efectuando las operaciones elementales entre filas $e_1, e_2, \ldots, e_q$ sobre la matriz identidad I.

La discusión anterior nos conduce al siguiente algoritmo (eliminación gaussiana) que bien halla la inversa de una matriz n-cuadrada A, o bien determina que A no es invertible.

Algoritmo 4.9: Inversa de una matriz A

Paso 1. Construir la matriz [por bloques] $n \times 2n$ $M = (A \mid I)$; esto es, A está en la mitad izquierda de M y la matriz identidad I en la derecha.

Paso 2. Reducir por filas M a forma escalonada. Si el proceso genera una fila nula en la mitad A de M, terminar (A no es invertible). En caso contrario, la mitad A adoptará forma triangular.

Paso 3. Más aún, reducir M a la forma canónica por filas $(I \mid B)$, donde I ha reemplazado a A en la mitad izquierda de la matriz.

Paso 4. Tomar $A^{-1} = B$.

EJEMPLO 4.12. Supongamos que queremos encontrar la inversa de $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$. Primero construimos la matriz por bloques $M = (A \mid I)$ y la reducimos a forma escalonada:

$$M = \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right)$$

En forma escalonada, la mitad izquierda de M está en forma triangular; por consiguiente, A es invertible. A continuación reducimos M a su forma canónica por filas:

$$M \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right)$$

La matriz identidad ocupa la mitad izquierda de la matriz final; de aquí, la mitad derecha es A^{-1}. Dicho de otro modo,

$$A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$$

4.10. OPERACIONES ELEMENTALES ENTRE COLUMNAS. EQUIVALENCIA DE MATRICES

Esta sección repite parte de la discusión de la precedente, utilizando las columnas de una matriz en vez de sus filas. (La elección de utilizar primero las filas proviene del hecho de que las operaciones entre filas están fuertemente ligadas a las operaciones con ecuaciones lineales.) Mostramos también la relación existente entre las operaciones entre filas y columnas y sus matrices elementales.

Las operaciones elementales entre columnas análogas a las operaciones elementales entre filas son las siguientes:

[F_1] (Intercambio de columnas) Intercambiar las columnas i-ésima y j-ésima:

$$C_i \leftrightarrow C_j$$

[F_2] (Cambio de escala de columnas) Multiplicar la columna i-ésima por un escalar no nulo k:

$$kC_i \rightarrow C_i \qquad (k \neq 0)$$

[F_3] (Adición de columnas) Sustituir la columna i-ésima por ella misma más k veces la j-ésima:

$$kC_j + C_i \rightarrow C_i$$

Cada una de las operaciones anteriores tiene una inversa del mismo tipo, tal y como sucedía con las operaciones entre filas correspondientes.

Denotemos por f una operación elemental entre columnas. La matriz F, obtenida efectuando f sobre la matriz identidad I, es decir,

$$F = f(I)$$

se conoce como la *matriz elemental* correspondiente a la operación elemental entre columnas f.

EJEMPLO 4.13. Las matrices 3-cuadradas elementales correspondientes a las operaciones elementales entre columnas $C_3 \to C_1$, $-2C_3 \to C_3$ y $-5C_2 + C_3 \to C_3$ son, respectivamente,

$$F_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad F_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad F_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix}$$

A lo largo de toda la discusión posterior, e y f denotarán, respectivamente, operaciones elementales entre filas y columnas correspondientes y E y F sus matrices elementales asociadas.

Lema 4.13: Supongamos que A es cualquier matriz. Entonces

$$f(A) = [e(A^T)]^T$$

esto es, efectuar la operación entre columnas f sobre una matriz A conduce al mismo resultado que efectuar la operación entre filas correspondientes e sobre A^T y luego tomar la traspuesta.

La demostración del lema se obtiene directamente del hecho de que las columnas de A son las filas de A^T, y viceversa.

El lema anterior muestra que

$$F = f(I) = [e(I^T)]^T = [e(I)]^T = E^T$$

En otras palabras,

Corolario 4.14: F es la traspuesta de E.
(De este modo, F es invertible porque E lo es.) Además, de acuerdo con el lema precedente.

$$f(A) = [e(A^T)]^T = [EA^T]^T = (A^T)^T E^T = AF$$

lo que prueba el siguiente teorema (que es análogo al Teorema 4.9 relativo a las operaciones elementales entre filas):

Teorema 4.15: Para cualquier matriz A, $f(A) = AF$.

Es decir, el resultado de efectuar una operación elemental entre columnas f sobre una matriz A puede obtenerse multiplicando por la derecha A por la matriz elemental correspondiente F.

Se dice que una matriz B es *equivalente por columnas* a otra A si B puede obtenerse de A

mediante una sucesión finita de operaciones elementales entre columnas. Usando el argumento análogo al dado para el Teorema 4.12 llegamos al:

Teorema 4.16: B es equivalente por columnas a A si y sólo si existe una matriz no singular Q tal que $B = AQ$.

EQUIVALENCIA DE MATRICES

Se dice que una matriz B es *equivalente* a otra A si B puede obtenerse de A mediante una sucesión finita de operaciones elementales entre filas y columnas. Alternativamente (Problema 4.23), B es equivalente a A si existen matrices no singulares P y Q tales que $B = PAQ$. Al igual que la equivalencia por filas o por columnas, la equivalencia de matrices es una relación de equivalencia.

El principal resultado de esta subsección, demostrado en el Problema 4.25, se enuncia a continuación:

Teorema 4.17: Toda matriz $m \times n$ A es equivalente a una única matriz por bloques de la forma

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

donde I_r es la matriz identidad $r \times r$. (El entero no negativo r se llama el *rango* de A.)

4.11. MATRICES SIMETRICAS CONGRUENTES. LEY DE INERCIA

Se dice que una matriz B es *congruente* a otra A si existe una matriz no singular (invertible) P tal que

$$B = P^T AP$$

Por el Problema 4.123, la congruencia es una relación de equivalencia.

Supongamos que A es simétrica, o sea, $A^T = A$. En tal caso,

$$B^T = (P^T AP)^T = P^T A^T P^{TT} = P^T AP = B$$

y por tanto B es simétrica también. Dado que las matrices diagonales son simétricas, se deduce que únicamente matrices simétricas son congruentes a matrices diagonales.

El siguiente teorema juega un importante papel en el álgebra lineal.

Teorema 4.18 (ley de inercia): Sea A una matriz real simétrica. Entonces existe una matriz no singular P tal que $B = P^T AP$ es diagonal. Aún más, todas las matrices diagonales B que cumplan la condición anterior tienen el mismo número $\mathbf{p}$ de entradas positivas y el mismo número $\mathbf{n}$ de entradas negativas.

El *rango* y la *signatura* de la matriz anterior A se denotan y definen, respectivamente, por

$$\text{rango } A = \mathbf{p} + \mathbf{n} \qquad \text{y} \qquad \text{sig } A = \mathbf{p} - \mathbf{n}$$

Estos están unívocamente definidos, según el Teorema 4.19. [La noción de rango se define en realidad para cualquier matriz (Sección 5.7), y la definición precedente concuerda con la general.]

ALGORITMO DE DIAGONALIZACION

El algoritmo expuesto a continuación diagonaliza (bajo congruencia) una matriz real simétrica $A = (a_{ij})$.

Algoritmo 4.11: Diagonalización bajo congruencia de una matriz simétrica

Paso 1. Construir la matriz [por bloques] $n \times 2n$ $M = (A \mid I)$; esto es, A es la mitad izquierda de M y la matriz identidad I, la derecha.

Paso 2. Examinar la entrada a_{11}.

Caso I: $a_{11} \neq 0$. Efectuar las operaciones entre filas $-a_{i1}R_1 + a_{11}R_i \to R_i$, $i = 2, \ldots, n$, y después las operaciones entre columnas correspondientes $-a_{i1}C_1 + a_{11}C_i \to C_i$ para reducir la matriz M a la forma

$$M = \left(\begin{array}{cc|cc} a_{11} & 0 & * & * \\ 0 & B & * & * \end{array}\right) \qquad [1]$$

Caso II: $a_{11} = 0$, pero $a_{ii} \neq 0$ para algún $i > 1$. Efectuar la operación entre filas $R_1 \leftrightarrow R_i$ y luego la operación entre columnas correspondientes $C_1 \leftrightarrow C_i$ para llevar a_{ii} a la primera posición diagonal. Esto reduce la matriz a la del Caso I.

Caso III. Todas las entradas diagonales cumplen $a_{ii} = 0$. Elegir i, j tales que $a_{ij} \neq 0$ y efectuar la operación entre filas $R_j + R_i \to R_i$ junto a la operación entre columnas correspondiente $C_j + C_i \to C_i$ para llevar $2a_{ij} \neq 0$ a la i-ésima posición diagonal. Con esto se reduce la matriz a la del Caso II.

En cada uno de los casos reducimos finalmente M a la forma [1] en la que B es una matriz simétrica de orden menor que el de A.

Nota: Las operaciones entre filas variarán las dos mitades de M, pero las operaciones entre columnas sólo cambiarán su mitad izquierda.

Paso 3. Repetir el Paso 2 con cada nueva matriz (despreciando las primeras fila y columna de la precedente) hasta que A esté diagonalizada, esto es, hasta que M se transforme en una $M' = (D, Q)$, donde D es diagonal.

Paso 4. Tomar $P = Q^T$. Entonces $D = P^T AP$.

La justificación del algoritmo anterior se da a continuación. Sean $e_1, e_2, \ldots, e_k$ todas las operaciones elementales entre filas en el algoritmo y $f_1, f_2, \ldots, f_k$ las correspondientes operaciones elementales entre columnas. Supongamos que E_i y F_i son las matrices elementales asociadas. Según el Corolario 4.14,

$$F_i = E_i^T$$

Por el algoritmo anterior,

$$Q = E_k \cdots E_2 E_1 I = E_k \cdots E_2 E_1$$

ya que la mitad derecha I de M sólo es alterada por las operaciones entre filas. Por otra parte, la mitad izquierda A de M es transformada por las operaciones tanto entre filas como entre columnas; por consiguiente,

$$\begin{aligned} D &= E_k \cdots E_2 E_1 A F_1 F_2 \cdots F_k = \\ &= (E_k \cdots E_2 E_1) A (E_k \cdots E_2 E_1)^T = \\ &= QAQ^T = P^T A P \end{aligned}$$

donde $P = Q^T$.

EJEMPLO 4.14. Supongamos $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$, una matriz simétrica. Para encontrar una matriz no singular P tal que $B = P^T AP$ sea diagonal, empezamos construyendo la matriz por bloques $(A \mid I)$

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right)$$

Efectuamos las operaciones $-2R_1 + R_2 \to R_2$ y $3R_1 + R_3 \to R_3$ para $(A \mid I)$ y después las correspondientes $-2C_1 + C_2 \to C_2$ y $3C_1 + C_3 \to C_3$ sobre A obteniendo

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right)$$

A continuación efectuamos $-2R_2 + R_3 \to R_3$ y entonces la correspondiente $-2C_2 + C_3 \to C_3$ para obtener

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right)$$

Se ha diagonalizado A. Tomemos

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y entonces} \quad B = P^T AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Nótese que B tiene $\mathbf{p} = 2$ entradas positivas y $\mathbf{n} = 1$ entrada negativa.

4.12. FORMAS CUADRATICAS

Una *forma cuadrática* q en las variables $x_1, x_2, \ldots, x_n$ es un polinomio

$$q(x_1, x_2, \ldots, x_n) = \sum_{i<j} c_{ij} x_i x_j \qquad [4.1]$$

(donde cada término es de grado dos). Se dice que la forma cuadrática está *diagonalizada* si

$$q(x_1, x_2, \ldots, x_n) = c_{11}x_1^2 + c_{22}x_2^2 + \cdots + c_{nn}x_n^2$$

esto es, si q carece de términos con *productos cruzados* $x_i x_j$ (con $i \neq j$).

La forma cuadrática [4.1] puede expresarse de modo único en la forma matricial

$$q(X) = X^T AX \qquad [4.2]$$

donde $X = (x_1, x_2, \ldots, x_n)^T$ y $A = (a_{ij})$ es una matriz simétrica. Las entradas de A pueden obtenerse a partir de [4.1] tomando

$$a_{ii} = c_{ii} \qquad \text{y} \qquad a_{ij} = a_{ji} = c_{ij}/2 \qquad (\text{para } i \neq j)$$

es decir, A tiene entrada diagonal a_{ii} igual al coeficiente de x_i^2 y entradas a_{ij} y a_{ji} iguales a la mitad del coeficiente de $x_i x_j$. De este modo,

$$q(X) = (x_1, \ldots, x_n)\begin{pmatrix} a_{11} & a_{22} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= \sum_{i,j} a_{ij}x_i x_j = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2\sum_{i<j} a_{ij}x_i x_j$$

La matriz simétrica anterior A se denomina la *representación matricial* de la forma cuadrática q. Aunque muchas matrices A en [4.2] conducirán a la misma forma cuadrática q, sólo una de ellas es simétrica.

Recíprocamente, toda matriz simétrica A define una forma cuadrática q mediante [4.2]. Siendo así, existe una correspondencia uno-a-uno entre las formas cuadráticas q y las matrices simétricas A. Aún más, una forma cuadrática q está diagonalizada si y sólo si la matriz simétrica asociada A es diagonal.

EJEMPLO 4.15

a) La forma cuadrática

$$q(x, y, z) = x^2 - 6xy + 8y^2 - 4xz + 5yz + 7z^2$$

puede expresarse en la forma matricial

$$q(x, y, z) = (x, y, z)\begin{pmatrix} 1 & -3 & -2 \\ -3 & 8 & \frac{5}{2} \\ -2 & \frac{5}{2} & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

donde la matriz definidora es simétrica. También puede expresarse en la forma matricial

$$q(x, y, z) = (x, y, z)\begin{pmatrix} 1 & -6 & -4 \\ 0 & 8 & 5 \\ 0 & 0 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

en la que la matriz definidora es triangular superior.

b) La matriz simétrica $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$ determina la forma cuadrática

$$q(x, y) = (x, y)\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 + 6xy + 5y^2$$

Nota: Por razones de índole teórica, supondremos siempre que una forma cuadrática q está representada por una matriz simétrica A. Dado que A se obtiene de q mediante división por 2, tendremos que suponer también que $1 + 1 \neq 0$ en nuestro cuerpo K. Esto es siempre cierto si K es el cuerpo real $\mathbf{R}$ o el complejo $\mathbf{C}$.

MATRIZ DE CAMBIO DE VARIABLES

Consideremos un cambio de variables, digamos de $x_1, x_2, \ldots, x_n$ a $y_1, y_2, \ldots, y_n$, por medio de una sustitución lineal invertible de la forma

$$x_i = p_{i1}y_1 + p_{i2}y_2 + \cdots + p_{in}y_n \qquad (i = 1, 2, \ldots, n)$$

(Aquí *invertible* significa que se puede despejar cada una de las variables y de forma única en términos de las x.) Tal sustitución lineal puede expresarse en la forma matricial

$$X = PY \tag{4.3}$$

donde

$$X = (x_1, x_2, \ldots, x_n)^T \qquad Y = (y_1, y_2, \ldots, y_n)^T \qquad \text{y} \qquad P = (p_{ij})$$

La matriz P recibe el nombre de *matriz de cambio de variables*; es no singular, puesto que la sustitución lineal es invertible.

Recíprocamente, cualquier matriz no singular P define una sustitución lineal invertible de variables, $X = PY$. Además,

$$Y = P^{-1}X$$

proporciona la fórmula que expresa las variables y en términos de las x.

Existe una interpretación geométrica de la matriz de cambio de variables P, que se ilustra en el próximo ejemplo.

EJEMPLO 4.16. Consideremos el plano cartesiano $\mathbf{R}^2$ con los ejes x e y usuales y la matriz no singular 2×2

$$P = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$$

Las columnas $u_1 = (2, 1)^T$ y $u_2 = (-1, 1)^T$ de P determinan un nuevo sistema de coordenadas en el plano, digamos con ejes s y t, como se muestra en la Figura 4-1. Esto es:

1. El eje s está en la dirección de u_1 y su unidad de longitud es la longitud de u_1.
2. El eje t está en la dirección de u_2 y su unidad de longitud es la longitud de u_2.

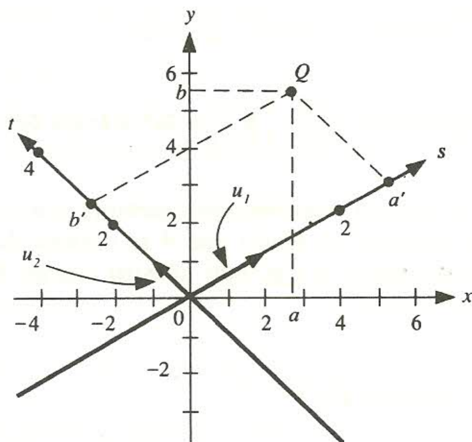


Figura 4-1.

Un punto cualquiera Q en el plano tendrá unas coordenadas respecto a cada sistema, digamos $Q(a, b)$ respecto a los ejes x e y y $Q(a', b')$ respecto a los s y t. Estos vectores de coordenadas están relacionados mediante la matriz P. Concretamente,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a' \\ b' \end{pmatrix} \quad \text{o} \quad X = PY$$

donde $X = (a, b)^T$ e $Y = (a', b')^T$.

DIAGONALIZACION DE UNA FORMA CUADRATICA

Consideremos una forma cuadrática q en las variables $x_1, x_2, \ldots, x_n$, es decir, $q(X) = X^T AX$ (donde A es una matriz simétrica). Supongamos que se efectúa un cambio de variables en q empleando la sustitución lineal [4.3]. Tomando $X = PY$ en q se llega a la forma cuadrática

$$q(Y) = (PY)^T A(PY) = Y^T(P^T AP)Y$$

Así $B = P^T AP$ es la representación matricial de la forma cuadrática en las nuevas variables $y_1, y_2, \ldots, y_n$. Observemos que la nueva matriz B es congruente a la inicial A, que representaba q.

Se dice que la sustitución lineal anterior $Y = PX$ *diagonaliza* la forma cuadrática $q(X)$ si $q(Y)$ es diagonal, esto es, si $B = P^T AP$ es una matriz diagonal. Como B es congruente a A y A es una matriz simétrica, podemos enunciar otra vez el Teorema 4.18 como sigue.

Teorema 4.19 (ley de inercia): Sea $q(X) = X^T AX$ una forma cuadrática real. Entonces existe una sustitución lineal invertible $Y = PX$ que diagonaliza q. Aún más, cualquier otra representación diagonal de q tiene el mismo número $\mathbf{p}$ de entradas positivas y el mismo número $\mathbf{n}$ de entradas negativas.

El *rango* y la *signatura* de una forma cuadrática q se denotan y definen por

$$\text{rango } q = \mathbf{p} + \mathbf{n} \qquad \text{y} \qquad \text{sig } q = \mathbf{p} - \mathbf{n}$$

Están unívocamente definidos, de acuerdo con el Teorema 4.19.

Dado que diagonalizar una forma cuadrática q equivale a diagonalizar bajo congruencia una matriz simétrica, el Algoritmo 4.11 puede utilizarse también aquí.

EJEMPLO 4.17. Consideremos la forma cuadrática

$$q(x, y, z) = x^2 + 4xy + 5y^2 - 6xz - 8yz + 8z^2 \qquad [1]$$

La matriz simétrica A que representa q es:

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$$

Según el Ejemplo 4.14, la matriz no singular P siguiente diagonaliza A bajo congruencia:

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad B = P^TAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

De acuerdo con esto, q puede diagonalizarse mediante la sustitución lineal:

$$\begin{aligned} x &= r - 2s + 7t \\ y &= s - 2t \\ z &= t \end{aligned}$$

De forma específica, sustituyendo x, y y z en [1] se llega a la forma cuadrática:

$$q(r, s, t) = r^2 + s^2 - 5t^2 \qquad [2]$$

Aquí $p = 2$ y $n = 1$; por tanto,

$$\text{rango } q = 3 \quad \text{y} \quad \text{sig } q = 1$$

Nota: Existe una interpretación geométrica de la ley de inercia (Teorema 4.19) que damos aquí utilizando la forma cuadrática del Ejemplo 4.17. Consideremos la siguiente superficie S en $\mathbf{R}^3$:

$$q(x, y, z) = x^2 + 4xy + 5y^2 - 6xz - 8yz + 8z^2 = 25$$

Bajo el cambio de variables

$$x = r - 2s + 7t \qquad y = s - 2t \qquad z = t$$

o, equivalentemente, relativa a un nuevo sistema de coordenadas con ejes r, s y t, la ecuación de S se convierte en

$$q(r, s, t) = r^2 + s^2 - 5t^2 = 25$$

De acuerdo con ello, S es un hiperboloide de una hoja, ya que hay dos entradas positivas y una negativa en la diagonal. Aún más, S será siempre un hiperboloide de una hoja, con independencia del sistema de coordenadas. Siendo así, cualquier representación diagonal de la forma cuadrática $q(x, y, z)$ contendrá dos entradas positivas y una negativa en la diagonal.

MATRICES SIMETRICAS Y FORMAS CUADRATICAS DEFINIDAS POSITIVAS

Se dice que una matriz real simétrica A es *definida positiva* si

$$X^T AX > 0$$

para todo vector (columna) no nulo X en $\mathbf{R}^n$. Análogamente, se dice que una forma cuadrática q es *definida positiva* si $q(v) > 0$ para todo vector no nulo en $\mathbf{R}^n$.

De forma alternativa, una matriz real simétrica A o su forma cuadrática q es *definida positiva* si cualquier representación diagonal tiene sólo entradas positivas. Tales matrices y formas cuadráticas desempeñan un papel muy importante dentro del álgebra lineal. Se estudiarán en los Problemas 4.54-4.60.

4.13. SIMILARIDAD

Una función $f: \mathbf{R}^n \to \mathbf{R}^n$ puede verse geométricamente como «enviando» o «aplicando» cada punto Q sobre un punto $f(Q)$ en el espacio $\mathbf{R}^n$. Supongamos que la función f puede representarse de la forma

$$f(Q) = AQ$$

donde A es una matriz $n \times n$ y las coordenadas de Q se escriben como un vector columna. Aún más, supongamos que P es una matriz no singular que puede verse como introduciendo un nuevo sistema de coordenadas en el espacio $\mathbf{R}^n$. (Véase Ejemplo 4.16.) Probamos que, referida a este nuevo sistema de coordenadas, f está representada por la matriz

$$B = P^{-1}AP$$

esto es,

$$f(Q') = BQ'$$

donde Q' denota el vector columna de las coordenadas de Q respecto al nuevo sistema de coordenadas.

EJEMPLO 4.18. Consideremos la función $f: \mathbf{R}^2 \to \mathbf{R}^2$ definida por

$$f(x, y) = (3x - 4y,\ 5x + 2y)$$

o, equivalentemente,

$$f\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{donde} \qquad A = \begin{pmatrix} 3 & -4 \\ 5 & 2 \end{pmatrix}$$

Supongamos que se introduce en $\mathbf{R}^2$ un nuevo sistema de coordenadas con ejes s y t por medio de la matriz no singular

$$P = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \qquad \text{y por tanto} \qquad P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

(Véase la Figura 4-1.) Respecto a este nuevo sistema de coordenadas de $\mathbf{R}^2$, la función f puede representarse como

$$f\begin{pmatrix} s \\ t \end{pmatrix} = B\begin{pmatrix} s \\ t \end{pmatrix}$$

donde

$$B = P^{-1}AP = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} 3 & -4 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{14}{3} & -\frac{10}{3} \\ \frac{22}{3} & \frac{1}{3} \end{pmatrix}$$

En otras palabras,

$$f(s, t) = (\tfrac{14}{3}s - \tfrac{10}{3}t,\ \tfrac{22}{3}s + \tfrac{1}{3}t)$$

La discusión precedente nos conduce a la siguiente:

Definición: Una matriz B es *similar* a otra A si existe una matriz no singular P tal que

$$B = P^{-1}AP$$

La similaridad, como la congruencia, es una relación de equivalencia (Problema 4.125); por consiguiente, podemos decir que A y B son matrices similares cuando $B = P^{-1}AP$.

Se dice que una matriz A es *diagonalizable* si existe una matriz no singular P tal que $B = P^{-1}AP$ es una matriz diagonal. La cuestión de si una matriz dada A es o no diagonalizable y el hallar la matriz P en caso de serlo juegan un papel importante en el álgebra lineal. Estas cuestiones se acometerán en el Capítulo 8.

4.14. FACTORIZACION LU

Supongamos que A es una matriz no singular que puede llevarse a forma triangular (superior) U mediante el uso de operaciones de adición de filas exclusivamente, esto es, supongamos que A puede triangularizarse por medio del siguiente algoritmo, que escribimos empleando la notación característica de los ordenadores.

Algoritmo 4.14. Triangularización de la matriz $A = (a_{ij})$

Paso 1. Repetir para $i = 1, 2, \ldots, n - 1$;

Paso 2. Repetir para $j = i + 1, \ldots, n$:

a) Tomar $m_{ij} := a_{ij}/a_{ii}$.

b) Tomar $R_j := m_{ij}R_i + R_j$.

[Fin del bucle interno Paso 2.]

[Fin del bucle externo Paso 1.]

Paso 3. Salir.

Los números m_{ij} se denominan *multiplicadores*. A veces seguimos la pista de dichos multiplicadores a través de la siguiente matriz triangular inferior L:

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -m_{21} & 1 & 0 & \dots & 0 & 0 \\ -m_{31} & -m_{32} & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -m_{n1} & -m_{n2} & -m_{n3} & \dots & -m_{n,n-1} & 1 \end{pmatrix}$$

O sea, L tiene unos en la diagonal, ceros sobre ésta y los opuestos de los m_{ij} como entradas ij bajo la misma.

La matriz triangular inferior precedente L puede describirse alternativamente como se hace a continuación. Denotemos por $e_1, e_2, \dots, e_k$ la sucesión de operaciones elementales ente filas en el algoritmo anterior. Las inversas de dichas operaciones son las siguientes. Para $i = 1, 2, \dots, n-1$ tenemos

$$-m_{ij}R_i + R_j \to R_j \qquad (j = i+1, \dots, n)$$

Efectuar estas operaciones en orden inverso sobre la matriz identidad I proporciona la matriz L. Así pues,

$$L = E_1^{-1}E_2^{-1} \cdots E_k^{-1}I$$

donde $E_1, \dots, E_k$ son las matrices elementales asociadas a las operaciones $e_1, \dots, e_k$.

Por otra parte, las operaciones elementales $e_1, \dots, e_k$ transforman la matriz original A en la matriz triangular superior U. Siendo así, $E_k \cdots E_2E_1A = U$. De acuerdo con esto,

$$A = (E_1^{-1}E_2^{-1} \cdots E_k^{-1})U = (E_1^{-1}E_2^{-1} \cdots E_k^{-1}I)U = LU$$

Esto nos da la factorización LU clásica de una matriz A. Establezcamos formalmente este resultado como un teorema.

Teorema 4.20: Sea A una matriz como la precedente. Entonces $A = LU$, donde L es una matriz triangular inferior con unos en la diagonal y U una matriz triangular superior sin ceros en la diagonal.

Nota: Subrayamos el hecho de que el teorema anterior sólo se aplica a matrices no singulares A que pueden llevarse a forma triangular sin ningún intercambio de filas. Se dice que tales matrices son *factorizables* LU o que tienen una factorización LU.

EJEMPLO 4.19. Sea $A = \begin{pmatrix} 1 & 2 & -3 \\ -3 & -4 & 13 \\ 2 & 1 & -5 \end{pmatrix}$. Entonces A puede reducirse a forma triangular por medio de las operaciones $3R_1 + R_2 \to R_2$ y $-2R_1 + R_3 \to R_3$, y luego $(\frac{3}{2})R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & -3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{pmatrix}$$

Esto nos da la factorización $A = LU$, donde

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{2} & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{pmatrix}$$

Adviértase que las entradas -3, 2 y $-\frac{3}{2}$ de L provienen de las operaciones elementales entre filas anteriores y que U es la forma triangular de A.

APLICACIONES A LAS ECUACIONES LINEALES

Consideremos un algoritmo de ordenador, M. Denotemos por $C(n)$ el tiempo de ejecución del algoritmo en función del tamaño n de la entrada de datos. [La función $C(n)$ se llama a veces la *complejidad temporal* o, simplemente, la *complejidad* del algoritmo.] Con frecuencia, $C(n)$ simplemente cuenta el número de multiplicaciones y divisiones efectuadas por M, pero no el número de adiciones y sustracciones debido a que se invierte mucho menos tiempo en su ejecución.

Consideremos ahora un sistema de ecuaciones lineales

$$AX = B$$

donde $A = (a_{ij})$ tiene factorización LU y

$$X = (x_1, \ldots, x_n)^T \qquad \text{y} \qquad B = (b_1, \ldots, b_n)^T$$

Entonces el sistema puede llevarse a forma triangular (para poder efectuar la sustitución hacia atrás) aplicando el algoritmo anterior a la matriz ampliada $M = (A, B)$ del mismo. La complejidad temporal del algoritmo y de la sustitución hacia atrás son, respectivamente,

$$C(n) \approx n^3/2 \qquad \text{y} \qquad C(n) \approx n^2/2$$

siendo n el número de ecuaciones.

Por otra parte, supongamos que disponemos ya de la factorización $A = LU$. En tal caso, para triangularizar el sistema sólo es necesario efectuar las operaciones entre filas del algoritmo (retenidas en la matriz L) sobre el vector columna B, con lo que la complejidad temporal es

$$C(n) \approx n^2/2$$

Por supuesto, obtener la factorización LU requiere el uso del algoritmo original, donde $C(n) \approx n^3/2$. De este modo, podría no ganarse nada encontrando primero la factorización LU cuando se halla implicado un solo sistema. No obstante, existen situaciones, ilustradas más adelante, en las que la factorización LU resulta útil.

Supongamos que para una matriz dada A necesitamos resolver el sistema

$$AX = B$$

repetidamente, para una sucesión de vectores constantes diferentes, digamos $B_1, B_2, \ldots, B_k$. Supongamos además que algunos de los B_i dependen de la solución del sistema obtenida al

utilizar los vectores precedentes B_j. En tal caso, resulta más eficaz hallar primero la factorización LU de A y después emplearla para resolver el sistema para cada nuevo B.

EJEMPLO 4.20. Consideremos el sistema

$$\begin{aligned} x - 2y - z &= k_1 \\ 2x - 5y - z &= k_2 \quad \text{o} \quad AX = B \\ -3x + 10y - 3z &= k_3 \end{aligned} \qquad [1]$$

donde $A = \begin{pmatrix} 1 & -2 & -1 \\ 2 & -5 & -1 \\ -3 & 10 & -3 \end{pmatrix}$ y $B = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$.

Supongamos que queremos resolverlo para B_1, B_2, B_3, B_4, donde $B_1 = (1, 2, 3)^T$ y

$$B_{j+1} = B_j + X_j \qquad (\text{para } j > 1)$$

siendo X_j la solución de [1] usando B_j. Aquí es más eficaz empezar por obtener la factorización LU de A y luego utilizarla para resolver el sistema con cada uno de los vectores B. (Véase el Problema 4.73.)

PROBLEMAS RESUELTOS

ALGEBRA DE MATRICES CUADRADAS

4.1. Sea $A = \begin{pmatrix} 1 & 3 & 6 \\ 2 & -5 & 8 \\ 4 & -2 & 7 \end{pmatrix}$. Determinar: *a*) la diagonal y la traza de A; *b*) $A(u)$, donde $u = (2, -3, 5)^T$; *c*) $A(v)$, siendo $v = (1, 7, -2)$.

a) La diagonal consiste en los elementos situados desde la esquina superior izquierda hasta la inferior derecha de la matriz, esto es, los elementos a_{11}, a_{22}, a_{33}. De este modo, la diagonal de A consiste en los escalares 1, -5 y 7. La traza de A es la suma de los elementos diagonales; por consiguiente, tr $A = 1 - 5 + 7 = 3$.

b) $$A(u) = Au = \begin{pmatrix} 1 & 3 & 6 \\ 2 & -5 & 8 \\ 4 & -2 & 7 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 - 9 + 30 \\ 4 + 15 + 40 \\ 8 + 6 + 35 \end{pmatrix} = \begin{pmatrix} 23 \\ 59 \\ 49 \end{pmatrix}$$

c) Según nuestro convenio, $A(v)$ no está definido para un vector fila v.

4.2. Sea $A = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}$. *a*) Hallar A^2 y A^3. *b*) Hallar $f(A)$, donde $f(x) = 2x^3 - 4x + 5$.

a)
$$A^2 = AA = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix} = \begin{pmatrix} 1+8 & 2-6 \\ 4-12 & 8+9 \end{pmatrix} = \begin{pmatrix} 9 & -4 \\ -8 & 17 \end{pmatrix}$$
$$A^3 = AA^2 = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 9 & -4 \\ -8 & 17 \end{pmatrix} = \begin{pmatrix} 9-16 & -4+34 \\ 36+24 & -16-51 \end{pmatrix} = \begin{pmatrix} -7 & 30 \\ 60 & -67 \end{pmatrix}$$

b) Para encontrar $f(A)$, primero sustituimos x por A y la constante 5 por $5I$ en el polinomio dado, $f(x) = 2x^3 - 4x + 5$:

$$f(A) = 2A^3 - 4A + 5I = 2\begin{pmatrix} -7 & 30 \\ 60 & -67 \end{pmatrix} - 4\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Después multiplicamos cada matriz por su respectivo escalar:

$$f(A) = \begin{pmatrix} -14 & 60 \\ 120 & -134 \end{pmatrix} + \begin{pmatrix} -4 & -8 \\ -16 & 12 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

Por último, sumamos los elementos correspondientes de las matrices:

$$f(A) = \begin{pmatrix} -14-4+5 & 60-8+0 \\ 120-16+0 & -134+12+5 \end{pmatrix} = \begin{pmatrix} -13 & 52 \\ 104 & -117 \end{pmatrix}$$

4.3. Sea $A = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}$. Hallar $g(A)$, siendo $g(x) = x^2 - x - 8$.

$$A^2 = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4+6 & 4-2 \\ 6-3 & 6+1 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix}$$

$$g(A) = A^2 - A - 8I = \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} - 8\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$
$$= \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} + \begin{pmatrix} -2 & -2 \\ -3 & 1 \end{pmatrix} + \begin{pmatrix} -8 & 0 \\ 0 & -8 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Así A es un cero de $g(x)$.

4.4. Se da la matriz $A = \begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}$. Encontrar un vector columna *no nulo* $u = \begin{pmatrix} x \\ y \end{pmatrix}$ tal que $A(u) = 3u$.

Comenzamos por establecer la ecuación matricial $Au = 3u$:

$$\begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix}$$

Escribimos cada miembro como una sola matriz (vector columna):

$$\begin{pmatrix} x+3y \\ 4x-3y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$$

Igualamos entre sí los elementos correspondientes para llegar a un sistema de ecuaciones y lo reducimos a forma escalonada:

$$\left.\begin{matrix} x + 3y = 3x \\ 4x - 3y = 3y \end{matrix}\right\} \to \left\{\begin{matrix} 2x - 3y = 0 \\ 4x - 6y = 0 \end{matrix}\right\} \to \left\{\begin{matrix} 2x - 3y = 0 \\ 0 = 0 \end{matrix}\right\} \to 2x - 3y = 0$$

El sistema se reduce a una ecuación homogénea con dos incógnitas y por tanto tiene un número infinito de soluciones. Para obtener una solución no nula, tomemos, por ejemplo, $y = 2$; entonces $x = 3$. Esto es, $u = (3, 2)^T$ tiene la propiedad deseada.

4.5. Sea $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ 5 & 12 & -5 \end{pmatrix}$. Hallar todos los vectores $u = (x, y, z)^T$ tales que $A(u) = 0$.

Establecemos la ecuación $Au = 0$ y luego escribimos cada miembro como una sola matriz:

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ 5 & 12 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} x + 2y - 3z \\ 2x + 5y - z \\ 5x + 12y - 5z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Igualamos entre sí los elementos correspondientes para conseguir un sistema homogéneo y reducimos éste a forma escalonada:

$$\left.\begin{matrix} x + 2y - 3z = 0 \\ 2x + 5y - z = 0 \\ 5x + 12y - 5z = 0 \end{matrix}\right\} \to \left\{\begin{matrix} x + 2y - 3z = 0 \\ y + 5z = 0 \\ 2y + 10z = 0 \end{matrix}\right\} \to \left\{\begin{matrix} x + 2y - 3z = 0 \\ y + 5z = 0 \end{matrix}\right.$$

En la forma escalonada, z es la variable libre. Para llegar a la solución general, tomamos $z = a$, siendo a un parámetro. La sustitución hacia atrás conduce a $y = -5a$ y después a $x = 13a$. Siendo así, $u = (13a, -5a, a)^T$ representa todos los vectores tales que $Au = 0$.

4.6. Demostrar que la colección **M** de todas las matrices 2×2 de la forma $\begin{pmatrix} s & t \\ t & s \end{pmatrix}$ es un álgebra conmutativa de matrices.

Claramente, **M** es no vacía. Si $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ y $B = \begin{pmatrix} c & d \\ d & c \end{pmatrix}$ pertenecen a **M**, entonces

$$A + B = \begin{pmatrix} a + c & d + b \\ b + d & a + c \end{pmatrix} \qquad kA = \begin{pmatrix} ka & kb \\ kb & ka \end{pmatrix} \qquad AB = \begin{pmatrix} ac + bd & ad + bc \\ bc + ad & bd + ac \end{pmatrix}$$

también pertenecen a **M**. Así **M** es un álgebra de matrices. Además,

$$BA = \begin{pmatrix} ca + db & cb + da \\ da + cb & db + ca \end{pmatrix}$$

De este modo, $AB = BA$ y en consecuencia **M** es un álgebra conmutativa de matrices.

4.7. Encontrar todas las matrices $M = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ que conmutan con $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Primero hallamos

$$AM = \begin{pmatrix} x+z & y+t \\ z & t \end{pmatrix} \quad \text{y} \quad MA = \begin{pmatrix} x & x+y \\ z & z+t \end{pmatrix}$$

Luego imponemos $AM = MA$ para llegar a las cuatro ecuaciones

$$x + z = x \qquad y + t = x + y \qquad z = z \qquad t = z + t$$

De la primera ecuación, o de la última, $z = 0$; de la segunda, $x = t$. En ese caso, M es cualquier matriz de la forma $\begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$.

4.8. Sea $e_i = (0, \ldots, 1, \ldots, 0)^T$, donde $i = 1, \ldots, n$, el vector (columna) en $\mathbf{R}^n$ con 1 en la posición i-ésima y 0 en cualquier otra, y sean A y B dos matrices $m \times n$.

a) Probar que Ae_i es la i-ésima columna de A.

b) Supóngase $Ae_i = Be_i$ para cada i. Probar que $A = B$.

c) Supóngase $Au = Bu$ para todo vector u en $\mathbf{R}^n$. Demostrar que $A = B$.

a) Sean $A = (a_{ij})$ y $Ae_i = (b_1, \ldots, b_n)^T$. Entonces

$$b_k = R_k e_i = (a_{k1}, \ldots, a_{kn})(0, \ldots, 1, \ldots, 0)^T = a_{ki}$$

donde R_k es la fila k-ésima de A. Así

$$Ae_i = (a_{1i}, a_{2i}, \ldots, a_{ni})^T$$

que es la columna i-ésima de A.

b) $Ae_i = Be_i$ significa que A y B tienen la misma columna i-ésima para todo i. Por tanto, $A = B$.

c) Si $Au = Bu$ para todo vector u en $\mathbf{R}^n$, necesariamente $Ae_i = Be_i$ para cada i, con lo que $A = B$.

4.9. Supóngase que A es una matriz $m \times n$. Mostrar que: *a*) $I_m A = A$, *b*) $AI_n = A$. (De este modo, $AI = IA = A$ cuando A es una matriz cuadrada.)

Utilizamos el hecho de que $I = (\delta_{ij})$, donde δ_{ij} es la delta de Kronecker (Ejemplo 4.3).

a) Supongamos $I_m A = (f_{ij})$. Entonces

$$f_{ij} = \sum_{k=1}^{m} \delta_{ik} a_{kj} = \delta_{ii} a_{ij} = a_{ij}$$

Siendo así, $I_m A = A$, ya que las entradas correspondientes coinciden.

b) Supongamos $AI_n = (g_{ij})$. Entonces

$$g_{ij} = \sum_{k=1}^{n} a_{ik} \delta_{kj} = a_{ij} \delta_{jj} = a_{ij}$$

De este modo, $AI_n = A$, puesto que las entradas correspondientes son iguales.

4.10. Demostrar el Teorema 4.1.

i) Sea $A + B = (c_{ij})$. Entonces $c_{ij} = a_{ij} + b_{ij}$, por lo que

$$\operatorname{tr}(A + B) = \sum_{k=1}^{n} c_{kk} = \sum_{k=1}^{n} (a_{kk} + b_{kk}) = \sum_{k=1}^{n} a_{kk} + \sum_{k=1}^{n} b_{kk} = \operatorname{tr} A + \operatorname{tr} B$$

ii) Sea $kA = (c_{ij})$. Entonces $c_{ij} = ka_{ij}$, y

$$\text{tr } kA = \sum^n ka_{jj} = k \sum^n a_{jj} = k \cdot \text{tr } A$$

iii) Sean $AB = (c_{ij})$ y $BA = (d_{ij})$. Entonces $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$ y $d_{ij} = \sum_{k=1}^n b_{ik}a_{kj}$, de donde

$$\text{tr } AB = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik}b_{ki} = \sum_{k=1}^n \sum_{i=1}^n b_{ki}a_{ik} = \sum_{k=1}^n d_{kk} = \text{tr } BA$$

4.11. Demostrar el Teorema 4.2.

Supongamos $f(x) = \sum_{i=1}^r a_i x^i$ y $g(x) = \sum_{j=1}^s b_j x^j$.

i) Podemos tomar $r = s = n$ sumando potencias de x con coeficientes nulos. En tal caso,

$$f(x) + g(x) = \sum_{i=1}^n (a_i + b_i)x^i$$

De aquí

$$(f+g)(A) = \sum_{i=1}^n (a_i + b_i)A^i = \sum_{i=1}^n a_i A^i + \sum_{i=1}^n b_i A^i = f(A) + g(A)$$

ii) Tenemos $f(x)g(x) = \sum_{i,j} a_i b_j x^{i+j}$. Entonces

$$f(A)g(A) = \left(\sum_i a_i A^i\right)\left(\sum_j b_j A^j\right) = \sum_{i,j} a_i b_j A^{i+j} = (fg)(A)$$

iii) Usando $f(x)g(x) = g(x)f(x)$ tenemos

$$f(A)g(A) = (fg)(A) = (gf)(A) = g(A)f(A)$$

4.12. Sea $D_k = kI$ la matriz escalar asociada al escalar k. Probar que: *a*) $D_k A = kA$, *b*) $BD_k = kB$, *c*) $D_k + D_{k'} = D_{k+k'}$, *d*) $D_k D_{k'} = D_{kk'}$.

a) $D_k A = (kI)A = k(IA) = kA$.

b) $BD_k = B(kI) = k(BI) = kB$.

c) $D_k + D_{k'} = kI + k'I = (k + k')I = D_{k+k'}$.

d) $D_k D_{k'} = (kI)(k'I) = kk'(II) = kk'I = D_{kk'}$.

MATRICES INVERTIBLES. INVERSAS

4.13. Hallar la inversa de $\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$.

Método 1. Buscamos escalares x, y, z y w para los que

$$\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} 3x+5z & 3y+5w \\ 2x+3z & 2y+3w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

o que satisfagan

$$\begin{cases} 3x+5z=1 \\ 2x+3z=0 \end{cases} \quad \text{y} \quad \begin{cases} 3y+5w=0 \\ 2y+3w=1 \end{cases}$$

La solución del primer sistema es $x = -3$, $z = 2$, y la del segundo $y = 5$, $w = -3$. En ese caso, la inversa de la matriz dada es $\begin{pmatrix} -3 & 5 \\ 2 & -3 \end{pmatrix}$.

Método 2. La fórmula general para la inversa de la matriz 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es

$$A^{-1} = \begin{pmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad \text{donde} \qquad |A| = ad - bc$$

Así, si $A = \begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$, primero hallamos $|A| = (3)(3) - (5)(2) = -1 \neq 0$. A continuación intercambiamos los elementos diagonales, tomamos los opuestos de los otros elementos y multiplicamos por $1/|A|$:

$$A^{-1} = -1\begin{pmatrix} 3 & -5 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 5 \\ 2 & -3 \end{pmatrix}$$

4.14. Calcular las inversas de *a*) $A = \begin{pmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{pmatrix}$ y *b*) $B = \begin{pmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{pmatrix}$.

a) Construimos la matriz por bloques $M = (A \mid I)$ y la reducimos por filas a forma escalonada:

$$M = \left(\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ -1 & -1 & 5 & 0 & 1 & 0 \\ 2 & 7 & -3 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 3 & 5 & -2 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -5 & -3 & 1 \end{array}\right)$$

La mitad izquierda de M está ahora en forma triangular; por consiguiente, A tiene inversa. Vamos más allá, reduciendo por filas M a su forma canónica por filas:

$$M \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -9 & -6 & 2 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -16 & -11 & 3 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right) = (I \mid A^{-1})$$

Así $A^{-1} = \begin{pmatrix} -16 & -11 & 3 \\ \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$.

b) Formamos la matriz por bloques $M = (B \mid I)$ y la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 1 & 5 & -1 & | & 0 & 1 & 0 \\ 3 & 13 & -6 & | & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 0 & 2 & 3 & | & -1 & 1 & 0 \\ 0 & 4 & 6 & | & -3 & 0 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 0 & 2 & 3 & | & -1 & 1 & 0 \\ 0 & 0 & 0 & | & -1 & -2 & 1 \end{pmatrix}$$

En forma escalonada, M tiene una fila nula en su mitad izquierda; es decir, B no es reducible por filas a forma triangular. De acuerdo con ello, B no es invertible.

4.15. Demostrar las siguientes proposiciones:

a) Si A y B son invertibles, entonces AB también lo es y $(AB)^{-1} = B^{-1}A^{-1}$.

b) Si $A_1, A_2, \ldots, A_n$ son invertibles, entonces $(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$.

c) A es invertible si y sólo si lo es A^T.

d) Las operaciones de inversión y trasposición conmutan: $(A^T)^{-1} = (A^{-1})^T$.

a) Tenemos

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Por tanto, $B^{-1}A^{-1}$ es la inversa de AB.

b) Por inducción en n y utilizando la parte *a*),

$$(A_1 \cdots A_{n-1}A_n)^{-1} = [(A_1 \cdots A_{n-1})A_n]^{-1} = A_n^{-1}(A_1 \cdots A_{n-1})^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

c) Si A es invertible, existe una matriz B tal que $AB = BA = I$. En tal caso,

$$(AB)^T = (BA)^T = I^T \qquad \text{por lo que} \qquad B^TA^T = A^TB^T = I$$

Por consiguiente, A^T es invertible, con inversa B^T. El recíproco deriva del hecho de que $(A^T)^T = A$.

d) Según la parte *c*), B^T es la inversa de A^T, esto es, $B^T = (A^T)^{-1}$. Pero $B = A^{-1}$; de aquí $(A^{-1})^T = (A^T)^{-1}$.

4.16. Probar que, si tiene una fila o columna nula, A no es invertible.

De acuerdo con el Problema 3.20, si A tiene una fila nula, AB también la tendrá. Así, si A fuera invertible, $AA^{-1} = I$ implicaría que I tiene una fila nula. Por esta razón, A no es invertible. Por otra parte, si A tiene una columna nula, A^T debe tener una fila nula, por lo que A^T no será invertible. De este modo, A tampoco es invertible.

MATRICES ELEMENTALES

4.17. Encontrar las matrices 3-cuadradas elementales E_1, E_2, E_3 que corresponden, respectivamente, a las operaciones entre filas $R_1 \leftrightarrow R_2$, $-7R_3 \to R_3$ y $-3R_1 + R_2 \to R_2$.

Efectuamos las operaciones sobre la matriz identidad $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ para obtener

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4.18. Demostrar el Teorema 4.9.

Sea R_i la fila i-ésima de A, lo que denotamos escribiendo $A = (R_1, \ldots, R_m)$. Si B es una matriz para la que AB está definida, se sigue directamente de la definición de producto matricial que $AB = (R_1 B, \ldots, R_m B)$. Tomamos además

$$e_i = (0, \ldots, 0, \widehat{1}, 0, \ldots, 0), \qquad \widehat{\ } = i$$

Aquí $\widehat{\ } = i$ significa que 1 es la componente i-ésima. Por el Problema 4.8 sabemos que $e_i A = R_i$. Hacemos notar también que $I = (e_1, \ldots, e_m)$ es la matriz identidad.

i) Sea e la operación elemental entre filas $R_i \leftrightarrow R_j$. Entonces, para $\widehat{\ } = i$ y $\widehat{\widehat{\ }} = j$,

$$E = e(I) = (e_1, \ldots, \widehat{e_j}, \ldots, \widehat{\widehat{e_i}}, \ldots, e_m)$$

y

$$e(A) = (R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m)$$

De este modo,

$$EA = (e_1 A, \ldots, \widehat{e_j A}, \ldots, \widehat{\widehat{e_i A}}, \ldots, e_m A) = (R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m) = e(A)$$

ii) Ahora sea e la operación elemental entre filas $kR_i \to R_i$, $k \neq 0$. En tal caso, para $\widehat{\ } = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_i}, \ldots, e_m) \qquad \text{y} \qquad e(A) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m)$$

Así

$$EA = (e_1 A, \ldots, \widehat{ke_i A}, \ldots, e_m A) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m) = e(A)$$

iii) Finalmente, sea e la operación elemental entre filas $kR_j + R_i \to R_i$. Entonces, para $\widehat{\ } = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_j + e_i}, \ldots, e_m) \qquad \text{y} \qquad e(A) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m)$$

Usando $(ke_j + e_i)A = k(e_j A) + e_i A + kR_j + R_i$ tenemos

$$EA = (e_1 A, \ldots, \widehat{(ke_j + e_i)A}, \ldots, e_m A) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m) = e(A)$$

Por tanto, el teorema queda demostrado.

4.19. Probar las siguientes aserciones:

a) Cada una de las operaciones elementales entre filas expuestas a continuación tiene una operación inversa del mismo tipo.

[E_1] Intercambiar las filas i-ésima y j-ésima: $R_i \leftrightarrow R_j$.

[E_2] Multiplicar la fila i-ésima por un escalar no nulo k: $kR_i \to R_i$, $k \neq 0$.

[E_3] Sustituir la fila i-ésima por ella misma más k veces la j-ésima: $kR_j + R_i \to R_i$.

b) Toda matriz elemental E es invertible y su inversa es una matriz elemental.

a) Se trata cada operación por separado.

1. Intercambiando dos veces el mismo par de filas obtenemos la matriz original; esto es, esta operación es su propia inversa.
2. Multiplicando la fila i-ésima por k y después por k^{-1}, o por k^{-1} y luego por k, obtenemos la matriz original. En otras palabras, las operaciones $kR_i \to R_i$ y $k^{-1}R_i \to R_i$ son inversas.
3. Efectuando la operación $kR_j + R_i \to R_i$ y a continuación la $-kR_j + R_i \to R_i$, o primero la $-kR_j + R_i \to R_i$ y después la $kR_j + R_i \to R_i$, tendremos la matriz original. Dicho de otro modo, las operaciones $kR_j + R_i \to R_i$ y $-kR_j + R_i \to R_i$ son inversas.

b) Sea E la matriz elemental asociada a la operación elemental entre filas e: $e(I) = E$. Sea e' la operación inversa de e y E' su matriz elemental correspondiente. Entonces

$$I = e'(e(I)) = e'(E) = E'E \qquad \text{e} \qquad I = e(e'(I)) = e(E') = EE'$$

Por consiguiente, E' es la inversa de E.

4.20. Demostrar el Teorema 4.10.

Supongamos que A es invertible y equivalente por filas a una matriz B, en forma canónica por filas. En tal caso, existen matrices elementales $E_1, E_2, \ldots, E_s$ tales que $E_s \cdots E_2E_1A = B$. Dado que tanto A como cada E_i son invertibles, necesariamente lo es B. Pero si $B \neq I$, B tiene una fila nula y por ende no es invertible. En consecuencia, $B = I$ y *a*) implica *b*).

Si se verifica *b*), existen matrices elementales $E_1, E_2, \ldots, E_s$ tales que $E_s \cdots E_2E_1A = I$ y por tanto $A = (E_s \cdots E_2E_1)^{-1} = E_1^{-1}E_2^{-1} \cdots E_s^{-1}$. Pero las E_i^{-1} son también matrices elementales. De este modo, *b*) implica *c*).

Si se cumple *c*), $A = E_1E_2 \cdots E_s$. Las E_i son matrices invertibles; de aquí que también lo sea su producto, A. Así *c*) implica *a*). De acuerdo con ello, el teorema queda demostrado.

4.21. Demostrar el Teorema 4.11.

Supongamos que A no es invertible, de forma que no es equivalente por filas a la matriz identidad, sino que lo es a una matriz con una fila nula. Dicho de otro modo, existen matrices elementales $E_1, \ldots, E_s$ tales que $E_s \cdots E_2E_1A$ tiene una fila nula. Por consiguiente, $E_s \cdots E_2E_1AB = E_s \cdots E_2E_1$, que es una matriz invertible, tiene asimismo una fila nula. Pero las matrices invertibles no pueden tener filas nulas; por tanto, A es invertible, con inversa A^{-1}. Además,

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}$$

4.22. Demostrar el Teorema 4.12.

Si $B \sim A$, entonces $B = e_s(\ldots(e_2(e_1(A)))\ldots) = E_s \cdots E_2E_1A = PA$, donde $P = E_s \cdots E_2E_1$ es no singular. Recíprocamente, supongamos que $B = PA$ con P no singular. Según el Teorema 4.10, P es un producto de matrices elementales y en consecuencia B puede obtenerse a partir de A mediante una sucesión de operaciones elementales entre filas, es decir, $B \sim A$. Queda así probado el teorema.

4.23. Probar que B es equivalente a A si y sólo si existen matrices invertibles P y Q tales que $B = PAQ$.

Si B es equivalente a A, entonces $B = E_s \cdots E_2E_1AF_1F_2 \cdots F_t \equiv PAQ$, donde $P = E_s \cdots E_2E_1$ y $Q = F_1F_2 \cdots F_t$ son invertibles. El recíproco deriva del hecho de que cada paso es reversible.

4.24. Probar que la equivalencia de matrices, escrita $\approx$, es una relación de equivalencia: *a*) $A \approx A$. *b*) Si $A \approx B$, entonces $B \approx A$. *c*) Si $A \approx B$ y $B \approx C$, entonces $A \approx C$.

a) $A = IAI$, donde I es no singular; de aquí $A \approx A$.

b) Si $A \approx B$, necesariamente $A = PBQ$, siendo P y Q no singulares. En ese caso, $B = P^{-1}AQ^{-1}$, donde P^{-1} y Q^{-1} son no singulares. Por consiguiente, $B \approx A$.

c) Si $A \approx B$ y $B \approx C$, entonces $A = PBQ$ y $B = P'CQ'$, donde P, Q, P', Q' son no singulares. En consecuencia,

$$A = P(P'CQ')Q = (PP')C(QQ')$$

siendo PP' y QQ' no singulares. De aquí se llega a $A \approx C$.

4.25. Demostrar el Teorema 4.17.

La demostración es constructiva, en forma de algoritmo.

Paso 1. Reducir A a forma canónica por filas, con entradas principales no nulas $a_{11}, a_{2j_2}, \ldots, a_{rj_r}$.

Paso 2. Intercambiar C_2 y C_{j_2}, C_3 y C_{j_3}, ..., y C_r y C_{j_r}. Esto proporciona una matriz de la forma $\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}$, con entradas principales no nulas $a_{11}, a_{22}, \ldots, a_{rr}$.

Paso 3. Utilizar operaciones entre columnas, con los a_{ii} como pivotes, para sustituir cada entrada de B por un cero; esto es, para

$$i = 1, 2, \ldots, r \qquad \text{y} \qquad j = r+1, r+2, \ldots, n$$

efectuar la operación $-b_{ij}C_i + C_j \to C_j$.

La matriz final tiene la forma deseada: $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

TIPOS ESPECIALES DE MATRICES

4.26. Encontrar una matriz triangular superior A tal que $A^3 = \begin{pmatrix} 8 & -57 \\ 0 & 27 \end{pmatrix}$.

Tomemos $A = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}$. Entonces A^3 tiene la forma $\begin{pmatrix} x^3 & * \\ 0 & z^3 \end{pmatrix}$. Así $x^3 = 8$, por lo que $x = 2$; $z^3 = 27$, por lo que $z = 3$. Seguidamente, calculamos A^3 usando $x = 2$ y $z = 3$:

$$A^2 = \begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 5y \\ 0 & 9 \end{pmatrix} \qquad \text{y} \qquad A^3 = \begin{pmatrix} 2 & y \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 4 & 5y \\ 0 & 9 \end{pmatrix} = \begin{pmatrix} 8 & 19y \\ 0 & 27 \end{pmatrix}$$

Por tanto, $19y = -57$, o $y = -3$. De acuerdo con esto, $A = \begin{pmatrix} 2 & -3 \\ 0 & 3 \end{pmatrix}$.

4.27. Demostrar el Teorema 4.3 iii).

Sea $AB = (c_{ij})$. Entonces

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \qquad \text{y} \qquad c_{ii} = \sum_{k=1}^{n} a_{ik}b_{ki}$$

Supongamos que $i > j$. En tal caso, para cualquier k, es $i > k$ o $k > j$, de modo que bien $a_{ik} = 0$ o bien $b_{kj} = 0$. Siendo así, $c_{ij} = 0$ y AB es triangular superior. Supongamos ahora que $i = j$. Entonces, para $k < i$, $a_{ik} = 0$; y para $k > i$, $b_{ki} = 0$. Por consiguiente, $c_{ii} = a_{ii}b_{ii}$, como se pretendía.

4.28. ¿Qué matrices son simultáneamente triangulares superiores e inferiores?

Si A es tanto triangular superior como triangular inferior, toda entrada fuera de la diagonal principal debe ser nula. Por tanto, A es diagonal.

4.29. Demostrar el Teorema 4.4.

i) $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$.

ii) $(A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T)$.

iii) Elegimos $B \equiv \frac{1}{2}(A + A^T)$ y $C \equiv \frac{1}{2}(A - A^T)$ y recurrimos a i) y ii).

Nótese que no hay otra elección posible.

4.30. Escribir $A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}$ como la suma de una matriz simétrica B y una antisimétrica C.

Calculamos $A^T = \begin{pmatrix} 2 & 7 \\ 3 & 8 \end{pmatrix}$, $A + A^T = \begin{pmatrix} 4 & 10 \\ 10 & 16 \end{pmatrix}$ y $A - A^T = \begin{pmatrix} 0 & -4 \\ 4 & 0 \end{pmatrix}$. Entonces

$$B = \frac{1}{2}(A + A^T) = \begin{pmatrix} 2 & 5 \\ 5 & 8 \end{pmatrix} \qquad C = \frac{1}{2}(A - A^T) = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$$

4.31. Determinar x, y, z, s, t si $A = \begin{pmatrix} x & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & y \\ z & s & t \end{pmatrix}$ es ortogonal.

Denotemos por R_1, R_2, R_3 las filas de A y por C_1, C_2, C_3 sus columnas. Como R_1 es un vector unitario, $x^2 + \frac{4}{9} + \frac{4}{9} = 1$ o $x = \pm\frac{1}{3}$. Como R_2 es un vector unitario, $\frac{4}{9} + \frac{1}{9} + y^2 = 1$ o $y = \pm\frac{2}{3}$. Dado que $R_1 \cdot R_2 = 0$, llegamos a $2x/3 + \frac{2}{9} + 2y/3 = 0$ o $3x + 3y = -1$. La única posibilidad es $x = \frac{1}{3}$ e $y = -\frac{2}{3}$. Así

$$A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ z & s & t \end{pmatrix}$$

Como las columnas son vectores unitarios,

$$\frac{1}{9} + \frac{4}{9} + z^2 = 1 \qquad \frac{4}{9} + \frac{1}{9} + s^2 = 1 \qquad \frac{4}{9} + \frac{4}{9} + t^2 = 1$$

Por tanto, $z = \pm\frac{2}{3}$, $s = \pm\frac{2}{3}$ y $t = \pm\frac{1}{3}$.

Caso i): $z = \frac{2}{3}$. Como C_1 y C_2 son ortogonales, $s = -\frac{2}{3}$; como C_1 y C_3 son ortogonales, $t = \frac{1}{3}$.

Caso ii): $z = -\frac{2}{3}$. Como C_1 y C_2 son ortogonales, $s = \frac{2}{3}$; como C_1 y C_3 son ortogonales, $t = -\frac{1}{3}$.

Por consiguiente, hay exactamente dos soluciones posibles:

$$A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

4.32. Supongamos que $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es ortogonal. Probar que $a^2 + b^2 = 1$ y

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad \text{o} \quad A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Dado que A es ortogonal, sus filas constituyen un conjunto ortonormal. De aquí

$$a^2 + b^2 = 1 \qquad c^2 + d^2 = 1 \qquad ac + bd = 0$$

Similarmente, las columnas forman un conjunto ortonormal, de modo que

$$a^2 + c^2 = 1 \qquad b^2 + d^2 = 1 \qquad ab + cd = 0$$

Por este motivo, $c^2 = 1 - a^2 = b^2$, de donde $c = \pm b$.

Caso i): $c = +b$. Entonces $b(a + d) = 0$ o $d = -a$; la matriz correspondiente es $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$.

Caso ii): $c = -b$. Entonces $b = (d - a) = 0$ o $d = a$; la matriz correspondiente es $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

4.33. Demostrar el Teorema 4.6.

Sean a y b cualesquiera números reales tales que $a^2 + b^2 = 1$. En ese caso, existe un número real θ tal que $a = \cos\theta$ y $b = \operatorname{sen}\theta$. El resultado se obtiene ahora del Problema 4.32.

4.34. Encontrar una matriz ortogonal 3×3, P, cuya primera fila sea un múltiplo de $u_1 = (1, 1, 1)$ y cuya segunda fila lo sea de $u_2 = (0, -1, 1)$.

Primero hallamos un vector u_3 ortogonal a u_1 y a u_2, como puede ser (el producto vectorial) $u_3 = u_1 \times u_2 = (2, -1, -1)$. Sea A la matriz con filas u_1, u_2 y u_3 y P la que se obtiene de A normalizando sus filas. De este modo,

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 2 & -1 & -1 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{pmatrix}$$

4.35. Demostrar el Teorema 4.7.

Supongamos $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Entonces

$$AA^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}$$

Dado que $AA^T = A^TA$, llegamos a

$$a^2 + b^2 = a^2 + c^2 \qquad c^2 + d^2 = b^2 + d^2 \qquad ac + bd = ab + cd$$

La primera ecuación conduce a $b^2 = c^2$; por tanto, $b = c$ o $b = -c$.

Caso i): $b = c$ (que incluye el caso $b = -c = 0$). Obtenemos la matriz simétrica $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$.

Caso ii): $b = -c \neq 0$. Entonces $ac + bd = b(d - a)$ y $ab + cd = b(a - d)$. Así $b(d - a) = b(a - d)$ y por tanto $2b(d - a) = 0$. Como $b \neq 0$, obtenemos $a = d$. Así A tiene la forma

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$$

que es la suma de una matriz escalar y una antisimétrica.

MATRICES COMPLEJAS

4.36. Determinar la conjugada de la matriz $A = \begin{pmatrix} 2+i & 3-5i & 4+8i \\ 6-i & 2-9i & 5+6i \end{pmatrix}$.

Tomamos el conjugado de cada elemento (siendo $\overline{a + bi} = a - bi$):

$$\bar{A} = \overline{\begin{pmatrix} 2+i & 3-5i & 4+8i \\ 6-i & 2-9i & 5+6i \end{pmatrix}} = \begin{pmatrix} 2-i & 3+5i & 4-8i \\ 6+i & 2+9i & 5-6i \end{pmatrix}$$

4.37. Hallar A^H cuando $A = \begin{pmatrix} 2-3i & 5+8i \\ -4 & 3-7i \\ -6-i & 5i \end{pmatrix}$.

$A^H = \bar{A}^T$, la traspuesta conjugada de A. Por tanto,

$$A^H = \begin{pmatrix} \overline{2-3i} & \overline{-4} & \overline{-6-i} \\ \overline{5+8i} & \overline{3-7i} & \overline{5i} \end{pmatrix} = \begin{pmatrix} 2+3i & -4 & -6+i \\ 5-8i & 3+7i & -5i \end{pmatrix}$$

4.38. Escribir $A = \begin{pmatrix} 2+6i & 5+3i \\ 9-i & 4-2i \end{pmatrix}$ en la forma $A = B + C$, donde B es hermítica y C antihermítica.

Comenzamos calculando

$$A^H = \begin{pmatrix} 2-6i & 9+i \\ 5-3i & 4+2i \end{pmatrix} \qquad A + A^H = \begin{pmatrix} 4 & 14+4i \\ 14-4i & 8 \end{pmatrix} \qquad A - A^H = \begin{pmatrix} 12i & -4+2i \\ 4+2i & -4i \end{pmatrix}$$

En consecuencia, las matrices requeridas son

$$B = \frac{1}{2}(A + A^H) = \begin{pmatrix} 2 & 7+2i \\ 7-2i & 4 \end{pmatrix} \quad \text{y} \quad C = \frac{1}{2}(A - A^H) = \begin{pmatrix} 6i & -2+i \\ 2+i & -2i \end{pmatrix}$$

4.39. Definir un conjunto ortonormal de vectores en $\mathbf{C}^n$ y demostrar el análogo complejo al Teorema 4.5:

Teorema: Sea A una matriz compleja. Las siguientes afirmaciones son equivalentes: *a*) A es unitaria. *b*) Las filas de A forman un conjunto ortonormal. *c*) Las columnas de A forman un conjunto ortonormal.

Los vectores $u_1, u_2, \ldots, u_r$ de $\mathbf{C}^n$ constituyen un conjunto ortonormal si $u_i \cdot u_j = \delta_{ij}$, donde el producto escalar en $\mathbf{C}^n$ se define según

$$(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}$$

y δ_{ij} es la delta de Kronecker. [Véase Ejemplo 4.3 *a*).]

Denotemos por $R_1, \ldots, R_n$ las filas de A; entonces $\bar{R}_1^T, \ldots, \bar{R}_n^T$ son las columnas de A^H. Sea $AA^H = (c_{ij})$. De acuerdo con la definición de producto matricial, $c_{ij} = R_i\bar{R}_j^T = R_i \cdot R_j$. En ese caso, $AA^H = I$ si y sólo si $R_i \cdot R_j = \delta_{ij}$, lo que se verifica si y sólo si $R_1, R_2, \ldots, R_n$ forman un conjunto ortonormal. Así *a*) y *b*) son equivalentes. De forma similar, A es unitaria si y sólo si A^H es unitaria, lo que se verifica si y sólo si las filas de A^H son ortonormales, para lo que es necesario y suficiente que lo sean las conjugadas de las columnas de A, lo que se da, finalmente, si y sólo si las columnas de A son ortonormales. De este modo, *a*) y *c*) son equivalentes y queda probado el teorema.

4.40. Mostrar que $A = \begin{pmatrix} \frac{1}{2} - \frac{2}{3}i & \frac{2}{3}i \\ -\frac{2}{3}i & -\frac{1}{3} - \frac{2}{3}i \end{pmatrix}$ es unitaria.

Las filas constituyen un conjunto ortonormal:

$$(\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) \cdot (\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) = (\tfrac{1}{9} + \tfrac{4}{9}) + \tfrac{4}{9} = 1$$

$$(\tfrac{1}{3} - \tfrac{2}{3}i, \tfrac{2}{3}i) \cdot (-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) = (\tfrac{2}{9}i + \tfrac{4}{9}) + (-\tfrac{2}{9}i - \tfrac{4}{9}) = 0$$

$$(-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) \cdot (-\tfrac{2}{3}i, -\tfrac{1}{3} - \tfrac{2}{3}i) = \tfrac{4}{9} + (\tfrac{1}{9} + \tfrac{4}{9}) = 1$$

Siendo así, A es unitaria.

MATRICES CUADRADAS POR BLOQUES

4.41. Determinar cuál de las siguientes matrices es cuadrada por bloques:

$$A = \left(\begin{array}{cc:c:cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ \hdashline 3 & 3 & 3 & 3 & 3 \\ 1 & 3 & 5 & 7 & 9 \end{array}\right) \qquad B = \left(\begin{array}{cc:c:cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ 3 & 3 & 3 & 3 & 3 \\ \hdashline 1 & 3 & 5 & 7 & 9 \end{array}\right)$$

Aunque A es una matriz cuadrada 5×5 y una matriz por bloques 3×3, los bloques diagonales segundo y tercero no son matrices cuadradas. Por consiguiente, A no es una matriz cuadrada por bloques

B es una matriz cuadrada por bloques.

4.42. Completar la partición de $C = \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ \hdashline 3 & 3 & 3 & 3 & 3 \\ 1 & 3 & 5 & 7 & 9 \end{array}\right)$ en una matriz cuadrada por bloques.

Una de las líneas horizontales está entre las filas segunda y tercera; por tanto, añadimos una línea vertical entre las columnas segunda y tercera. La otra línea horizontal está entre las filas cuarta y quinta; por consiguiente, añadimos una línea vertical entre las columnas cuarta y quinta. [Las

líneas horizontales y verticales deben situarse simétricamente para obtener una matriz cuadrada por bloques.] Esto nos conduce a la matriz cuadrada por bloques

$$C = \left(\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 3 & 3 & 3 & 3 & 3 \\ 1 & 3 & 5 & 7 & 9 \end{array}\right)$$

4.43. Determinar cuáles de las matrices cuadradas por bloques siguientes son triangulares inferiores, triangulares superiores o diagonales:

$$A = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 0 & 6 \end{array}\right) \qquad B = \left(\begin{array}{c|cc|c} 1 & 0 & 0 & 0 \\ \hline 2 & 3 & 4 & 0 \\ 5 & 0 & 6 & 0 \\ \hline 0 & 7 & 8 & 9 \end{array}\right) \qquad C = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 3 \\ 0 & 4 & 5 \end{array}\right). \qquad D = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 5 \\ \hline 0 & 6 & 7 \end{array}\right)$$

A es triangular superior porque el bloque bajo la diagonal es nulo.

B es triangular inferior porque todos los bloques sobre la diagonal son nulos.

C es diagonal porque los bloques sobre y bajo la diagonal son nulos.

D no es triangular superior ni inferior. Aún más, ninguna otra partición de D hará de ella una matriz triangular superior o inferior por bloques.

4.44. Considérense las siguientes matrices diagonales por bloques, de las cuales los bloques diagonales correspondientes tienen el mismo tamaño:

$$M = \text{diag}\,(A_1, A_2, \ldots, A_r) \qquad \text{y} \qquad N = \text{diag}\,(B_1, B_2, \ldots, B_r)$$

Hallar: *a*) $M + N$, *b*) kM, *c*) MN, *d*) $f(M)$ para un polinomio dado $f(x)$.

a) Nos limitamos a sumar los bloques diagonales: $M + N = \text{diag}\,(A_1 + B_1, A_2 + B_2, \ldots, A_r + B_r)$.

b) Simplemente multiplicamos los bloques diagonales por k: $kM = \text{diag}\,(kA_1, kA_2, \ldots, kA_r)$.

c) Simplemente multiplicamos los bloques diagonales correspondientes:

$$MN = \text{diag}\,(A_1B_1, A_2B_2, \ldots, A_rB_r)$$

d) Determinamos $f(A_i)$ para cada bloque diagonal A_i. Entonces $f(M) = \text{diag}\,(f(A_1), f(A_2), \ldots, f(A_r))$.

4.45. Encontrar M^2, siendo $M = \left(\begin{array}{cc|c|cc} 1 & 2 & & & \\ 3 & 4 & & & \\ \hline & & 5 & & \\ \hline & & & 1 & 3 \\ & & & 5 & 7 \end{array}\right)$.

Dado que M es diagonal por bloques, elevamos al cuadrado cada bloque:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

$$(5)(5) = (25)$$

$$\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 16 & 24 \\ 40 & 64 \end{pmatrix}$$

$$\text{Entonces } M^2 = \begin{pmatrix} 7 & 10 & & & \\ 15 & 22 & & & \\ & & 25 & & \\ & & & 16 & 24 \\ & & & 40 & 64 \end{pmatrix}.$$

MATRICES SIMETRICAS CONGRUENTES. FORMAS CUADRATICAS

4.46. Sea $A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$ una matriz simétrica. Hallar: *a*) una matriz no singular P tal que $P^T AP$ sea diagonal, esto es, la matriz diagonal $B = P^T AP$, *b*) la signatura de A.

a) Construimos primero la matriz por bloques $(A \mid I)$:

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right)$$

Efectuamos las operaciones entre filas $3R_1 + R_2 \to R_2$ y $-2R_1 + R_3 \to R_3$ sobre $(A \mid I)$ y luego las correspondientes operaciones entre columnas $3C_1 + C_2 \to C_2$ y $-2C_1 + C_3 \to C_3$ sobre A para obtener

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \quad \text{y después} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

A continuación efectuamos la operación entre filas $R_2 + 2R_3 \to R_3$, seguida de la correspondiente operación entre columnas $C_2 + 2C_3 \to C_3$ obteniendo

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \quad \text{y después} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

Se ha diagonalizado A. Tomamos

$$P = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{y entonces} \quad B = P^T AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix}$$

b) B tiene $\mathbf{p} = 2$ elementos diagonales positivos y $\mathbf{n} = 1$ negativo, de donde sig $A = 2 - 1 = 1$.

FORMAS CUADRATICAS

4.47. Determinar la forma cuadrática $q(x, y)$ correspondiente a la matriz simétrica $A = \begin{pmatrix} 5 & -3 \\ -3 & 8 \end{pmatrix}$.

$$q(x, y) = (x, y)\begin{pmatrix} 5 & -3 \\ -3 & 8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (5x - 3y, \ -3x + 8y)\begin{pmatrix} x \\ y \end{pmatrix} =$$
$$= 5x^2 - 3xy - 3xy + 8y^2 = 5x^2 - 6xy + 8y^2$$

4.48. Hallar la matriz simétrica A que corresponde a la forma cuadrática

$$q(x, y, z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$$

La matriz simétrica $A = (a_{ij})$ que representa $q(x_1, \ldots, x_n)$ tiene la entrada diagonal a_{ii} igual al coeficiente de x_i^2 y las entradas a_{ij} y a_{ji} iguales a la mitad del coeficiente de $x_i x_j$. Por tanto,

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & -1 & -3 \\ 4 & -3 & 1 \end{pmatrix}$$

4.49. Encontrar la matriz simétrica B, asociada a la forma cuadrática:

$a)$ $q(x, y) = 4x^2 + 5xy - 7y^2$ y $b)$ $q(x, y, z) = 4xy + 5y^2$

a) Aquí $B = \begin{pmatrix} 4 & \frac{5}{2} \\ \frac{5}{2} & -7 \end{pmatrix}$. (La división por 2 puede introducir fracciones incluso aunque los coeficientes en q sean enteros.)

b) A pesar de que sólo x e y aparecen en el polinomio, la expresión $q(x, y, z)$ indica que hay tres variables. En otras palabras,

$$q(x, y, z) = 0x^2 + 4xy + 5y^2 + 0xz + 0yz + 0z^2$$

Así

$$B = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.50. Considérense la forma cuadrática $q(x, y) = 3x^2 + 2xy - y^2$ y la sustitución lineal $x = s - 3t$, $y = 2s + t$.

a) Reescribir $q(x, y)$ en notación matricial y hallar la matriz A que representa la forma cuadrática.

b) Reescribir la sustitución lineal empleando notación matricial y encontrar la matriz P correspondiente a la sustitución.

c) Hallar $q(s, t)$ utilizando sustitución directa.

d) Hallar $q(s, t)$ usando notación matricial.

a) Aquí $q(x, y) = (x, y)\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$. Por tanto, $A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ y $q(X) = X^T AX$, donde $X = (x, y)^T$.

b) Tenemos $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix}$. De este modo, $P = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}$ y $X = PY$, donde $X = (x, y)^T$ e $Y = (s, t)^T$.

c) Sustituimos x e y en q obteniendo

$$\begin{aligned} q(s, t) &= 3(s - 3t)^2 + 2(s - 3t)(2s + t) - (2s + t)^2 = \\ &= 3(s^2 - 6st + 9t^2) + 2(2s^2 - 5st - 3t^2) - (s^2 + 4st + t^2) = 3s^2 - 32st + 20t^2 \end{aligned}$$

d) Aquí $q(X) = X^T AX$ y $X = PY$. Siendo así, $X^T = Y^T P^T$. Por consiguiente,

$$q(s, t) = q(Y) = Y^T P^T APY = (s, t)\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} =$$

$$= (s, t)\begin{pmatrix} 3 & -16 \\ -16 & 20 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = 3s^2 - 32st + 20t^2$$

[Tal y como cabía esperar, los resultados en *c*) y *d*) son iguales.]

4.51. Sea L una sustitución lineal $X = PY$, como en el Problema 4.50.

a) ¿Cuándo es L no singular?, ¿y ortogonal?

b) Describir una ventaja esencial de una sustitución ortogonal sobre una no singular.

c) ¿Es la sustitución lineal del Problema 4.50 no singular?, ¿y ortogonal?

a) Se dice que L es singular u ortogonal según lo sea la matriz P que la representa.

b) Recordemos que las columnas de la matriz P, que representa la sustitución lineal, introducen un nuevo sistema de referencia. Si P es ortogonal, los nuevos ejes son perpendiculares y tienen las mismas unidades de longitud que los originales.

c) La matriz $P = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}$ es no singular, pero no es ortogonal; en consecuencia, la sustitución lineal es no singular, pero no es ortogonal.

4.52. Sea $q(x, y, z) = x^2 + 4xy + 3y^2 - 8xz - 12yz + 9z^2$. Encontrar una sustitución lineal no singular que exprese las variables x, y, z en términos de las r, s, t, de forma que $q(r, s, t)$ sea diagonal. Asimismo, hallar la signatura de q.

Formamos la matriz por bloques $(A \mid I)$, siendo A la matriz que representa la forma cuadrática:

$$(A \mid I) = \left(\begin{array}{rrr|rrr} 1 & 2 & -4 & 1 & 0 & 0 \\ 2 & 3 & -6 & 0 & 1 & 0 \\ -4 & -6 & 9 & 0 & 0 & 1 \end{array}\right)$$

Efectuamos $-2R_1 + R_2 \to R_2$ y $4R_1 + R_4 \to R_3$ y las correspondientes operaciones entre columnas y a continuación $2R_2 + R_3 \to R_3$ y la operación entre columnas asociada para llegar a

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -7 & 4 & 0 & 1 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -3 & 0 & 2 & 1 \end{array}\right)$$

De este modo, la sustitución lineal $x = r - 2s$, $y = s + 2t$, $z = t$ proporcionará la forma cuadrática

$$q(r, s, t) = r^2 - s^2 - 3t^2$$

Además, sig $q = 1 - 2 = -1$.

4.53. Diagonalizar la siguiente forma cuadrática q mediante el método conocido como «completar el cuadrado»:

$$q(x, y) = 2x^2 - 12xy + 5y^2$$

Comenzamos sacando como factor común de los términos en x^2 y xy el coeficiente de x^2:

$$q(x, y) = 2(x^2 - 6xy) + 5y^2$$

Acto seguido, completamos el cuadrado entre paréntesis sumando un múltiplo apropiado de y^2, restando después la cantidad correspondiente fuera de los paréntesis. Con esto conseguimos

$$q(x, y) = 2(x^2 - 6xy + 9y^2) + 5y^2 - 18y^2 = 2(x - 3y)^2 - 13y^2$$

(El -18 proviene del hecho de que el $9y^2$ dentro de los paréntesis se multiplica por 2.) Sean $s = x - 3y$, $t = y$. En tal caso, $x = s + 3t$, $y = t$. La sustitución lineal conduce a la forma cuadrática $q(s, t) = 2s^2 - 13t^2$.

FORMAS CUADRATICAS DEFINIDAS POSITIVAS

4.54. Sea $q(x, y, z) = x^2 + 2y^2 - 4xz - 4yz + 7z^2$. ¿Es q definida positiva?

Diagonalizamos (bajo congruencia) la matriz simétrica A asociada a q (efectuando $2R_1 + R_3 \to R_3$ y $2C_1 + C_3 \to C_3$ y luego $R_2 + R_3 \to R_3$ y $C_2 + C_3 \to C_3$):

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La representación diagonal de q sólo contiene entradas positivas, 1, 2 y 1, en la diagonal, luego q es definida positiva.

4.55. Sea $q(x, y, z) = x^2 + y^2 + 2xz + 4yz + 3z^2$. ¿Es q definida positiva?

Diagonalizamos (bajo congruencia) la matriz simétrica A correspondiente a q:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Hay una entrada negativa, -2, en la representación diagonal de q, por lo que q no es definida positiva.

4.56. Probar que $q(x, y) = ax^2 + bxy + cy^2$ es definida positiva si y sólo si el discriminante $D = b^2 - 4ac < 0$.

Supongamos $v = (x, y) \neq 0$, digamos $y \neq 0$. Sea $t = x/y$. Entonces

$$q(v) = y^2[a(x/y)^2 + b(x/y) + c] = y^2(at^2 + bt + c)$$

De cualquier modo, $s = at^2 + bt + c$ está por encima del eje x, es decir, es positivo para cualquier valor de t si y sólo si $D = b^2 - 4ac < 0$. Siendo así, q es definida positiva si y sólo si $D < 0$.

4.57. Determinar qué forma cuadrática es definida positiva:

a) $q(x, y) = x^2 - 4xy + 5y^2$ *b)* $q(x, y) = x^2 + 6xy + 3y^2$

a) **Método 1.** Diagonalizamos completando el cuadrado:

$$q(x, y) = x^2 - 4xy + 4y^2 + 5y^2 - 4y^2 = (x - 2y)^2 + y^2 = s^2 + t^2$$

donde $s = x - 2y$, $t = y$. Así q es definida positiva.

Método 2. Calculamos el discriminante $D = b^2 - 4ac = 16 - 20 = -4$. Siendo $D < 0$, q es definida positiva.

b) **Método 1.** Diagonalizamos completando el cuadrado:

$$q(x, y) = x^2 + 6xy + 9y^2 + 3y^2 - 9y^2 = (x + 3y)^2 - 6y^2 = s^2 - 6t^2$$

donde $s = x + 3y$, $t = y$. Como -6 es negativo, q no es definida positiva.

Método 2. Calculamos $D = b^2 - 4ac = 36 - 12 = 24$. Siendo $D > 0$, no es definida positiva.

4.58. Sea B una matriz no singular arbitraria y sea $M = B^T B$. Demostrar que: *a*) M es simétrica, *b*) M es definida positiva.

a) $M^T = (B^T B)^T = B^T B^{TT} = B^T B = M$; por consiguiente, M es simétrica.

b) Dado que B es no singular, $BX \neq 0$ para todo $X \in \mathbf{R}^n$ no nulo. En consecuencia, el producto escalar de BX por él mismo, $BX \cdot BX = (BX)^T(BX)$, es positivo. Así

$$q(X) = X^T MX = X^T(B^T B)X = (X^T B^T)(BX) = (BX)^T(BX) > 0$$

De este modo, M es definida positiva.

4.59. Demostrar que $q(X) = \|X\|^2$, el cuadrado de la norma de un vector X, es una forma cuadrática definida positiva.

Para $X = (x_1, x_2, \ldots, x_n)$ tenemos $q(X) = x_1^2 + x_2^2 + \cdots + x_n^2$. Ahora bien, q es un polinomio con todos los términos de grado dos y está en forma diagonal con todas sus entradas positivas. Por tanto, q es una forma cuadrática definida positiva.

4.60. Demostrar que las dos definiciones siguientes de forma cuadrática definida positiva son equivalentes:

a) Las entradas diagonales son todas positivas en cualquier representación diagonal de q.

b) $q(Y) > 0$ para todo vector no nulo Y en $\mathbf{R}^n$.

Supongamos $q(Y) = a_1 y_1^2 + a_2 y_2^2 + \cdots + a_n x_n^2$. Si todos los coeficientes a_i son positivos, claramente $q(Y) > 0$ para todo vector no nulo Y en $\mathbf{R}^n$, de modo que *a*) implica *b*). Recíprocamente, supongamos $a_k < 0$. Sea $e_k = (0, \ldots, 1, \ldots, 0)$ el vector cuyas entradas son todas 0 exceptuando un 1 en la k-ésima posición. En tal caso, $q(e_k) = a_k < 0$ para $e_k \neq 0$. Así *b*) implica *a*). De acuerdo con esto, *a*) y *b*) son equivalentes.

SIMILARIDAD DE MATRICES

4.61. Considérese el plano $\mathbf{R}^2$ con los ejes usuales x e y. La matriz 2×2 no singular

$$P = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$$

determina un nuevo sistema de coordenadas en el plano, digamos con ejes s y t. (Véase Ejemplo 4.16.)

a) Trazar los nuevos ejes s y t en el plano $\mathbf{R}^2$.

b) Hallar las coordenadas de $Q(1, 5)$ en el nuevo sistema.

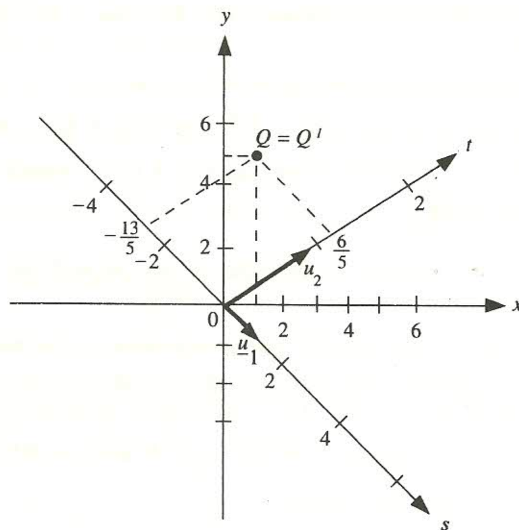


Figura 4-2.

a) Trazamos el eje s en la dirección de la primera columna $u_1 = (1, -1)^T$ de P con unidad de longitud igual a la longitud de u_1. De forma similar, trazamos el eje t en la dirección de la segunda columna $u_2 = (3, 2)^T$ de P con unidad de longitud igual a la longitud de u_2. Véase la Figura 4-2.

b) Encontramos $P^{-1} = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}$ usando, por ejemplo, la fórmula para la inversa de una matriz 2×2. Luego multiplicamos el vector (columna) de coordenadas de Q por P^{-1}:

$$P^{-1}Q = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} -\frac{13}{5} \\ \frac{6}{5} \end{pmatrix}$$

Así $Q'(-\frac{13}{5}, \frac{6}{5})$ representa Q en el nuevo sistema

4.62. Defínase $f: \mathbf{R}^2 \to \mathbf{R}^2$ según $f(x, y) = (2x - 5y, 3x + 4y)$.

a) Utilizando $X = (x, y)^T$, escribir f en notación matricial, esto es, hallar la matriz A tal que $f(X) = AX$.

b) Refiriéndose a los nuevos ejes coordenados s y t de $\mathbf{R}^2$ introducidos en el Problema 4.61 y usando $Y = (s, t)^T$, hallar $f(s, t)$ encontrando primero la matriz B tal que $f(Y) = BY$.

a) Aquí $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$; por consiguiente, $A = \begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}$.

b) Calculamos $B = P^{-1}AP = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{17}{5} & -\frac{59}{5} \\ \frac{6}{5} & \frac{13}{5} \end{pmatrix}$. Entonces

$$f\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} \frac{17}{5} & -\frac{59}{5} \\ \frac{6}{5} & \frac{13}{5} \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix}$$

Así $f(s, t) = (\frac{17}{5}s - \frac{59}{5}t, \frac{6}{5}s + \frac{13}{5}t)$.

4.63. Considérese el espacio $\mathbf{R}^3$ con los ejes x, y, z usuales. La matriz 3×3 no singular

$$P = \begin{pmatrix} 1 & 3 & -2 \\ -2 & -5 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

determina un nuevo sistema de coordenadas para $\mathbf{R}^3$, digamos con ejes r, s, t. [Alternativamente, P define la sustitución lineal $X = PY$, donde $X = (x, y, z)^T$ e $Y = (r, s, t)^T$.] Hallar las coordenadas del punto $Q(1, 2, 3)$ en el nuevo sistema.

Empezamos calculando P^{-1}. Construimos la matriz por bloques $M = (P \mid I)$ y la reducimos a forma canónica por filas:

$$M = \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ -2 & -5 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & -1 & 3 & -1 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 3 & 2 & 2 \\ 0 & 1 & 0 & 4 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -9 & -7 & -4 \\ 0 & 1 & 0 & 4 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)$$

De acuerdo con esto,

$$P^{-1} = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{y} \quad P^{-1}Q = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -47 \\ 16 \\ 6 \end{pmatrix}$$

Siendo así, $Q'(-47, 16, 6)$ representa Q en el nuevo sistema.

4.64. Defínase $f: \mathbf{R}^3 \to \mathbf{R}^3$ según

$$f(x, y, z) = (x + 2y - 3z, 2x + z, x - 3y + z)$$

y sea P la matriz no singular de cambio de variables del Problema 4.63. [De este modo, $X = PY$, siendo $X = (x, y, z)^T$ e $Y = (r, s, t)^T$.] Encontrar: *a*) la matriz A tal que $f(X) = AX$, *b*) la matriz B tal que $f(Y) = BY$, *c*) $f(r, s, t)$.

a) Los coeficientes de x, y y z proporcionan la matriz A:

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{y por tanto} \quad A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}$$

b) B es similar a A con respecto a P, esto es,

$$B = P^{-1}AP = \begin{pmatrix} -9 & -7 & -4 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 1 \\ 1 & -3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & -2 \\ -2 & -5 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -19 & 58 \\ 1 & 12 & -27 \\ 5 & 15 & -11 \end{pmatrix}$$

c) Utilizamos la matriz B para llegar a

$$f(r, s, t) = (r - 19s + 58t, r + 12s - 27t, 5r + 15s - 11t)$$

4.65. Supóngase que B es similar a A. Mostrar que tr $B =$ tr A.

Siendo B similar a A, existe una matriz no singular P tal que $B = P^{-1}AP$. Utilizando el Teorema 4.1,

$$\text{tr } B = \text{tr } P^{-1}AP = \text{tr } PP^{-1}A = \text{tr } A$$

FACTORIZACION *LU*

4.66. Encontrar la factorización LU de $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 5 & 6 \\ -3 & -2 & 7 \end{pmatrix}$.

Reducimos A a forma triangular por medio de las operaciones $-2R_1 + R_2 \to R_2$ y $3R_1 + R_3 \to R_3$ y luego $7R_2 + R_3 \to R_3$:

$$A \sim \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 7 & 13 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 27 \end{pmatrix}$$

Empleamos los opuestos de los multiplicadores -2, 3 y 7 de las operaciones entre filas precedentes para construir la matriz L, y la forma triangular de A para conseguir la matriz U; esto es,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -7 & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 27 \end{pmatrix}$$

(Como comprobación, multiplíquense L y U para verificar que $A = LU$.)

4.67. Hallar la factorización LDU de la matriz A del Problema 4.66.

La factorización $A = LDU$ se refiere a la situación en la que L es una matriz triangular inferior con unos en la diagonal (como en la factorización LU de A), D una matriz diagonal y U una matriz triangular superior con unos en la diagonal. De este modo, bastará sacar como factores las entradas diagonales de la matriz U en la factorización LU anterior para obtener las matrices D y U. Por consiguiente,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -7 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{27} \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

4.68. Hallar la factorización LU de $B = \begin{pmatrix} 1 & 4 & -3 \\ 2 & 8 & 1 \\ -5 & -9 & 7 \end{pmatrix}$.

Reducimos B a forma triangular efectuando primero las operaciones $-2R_1 + R_2 \rightarrow R_2$ y $5R_1 + R_3 \rightarrow R_3$:

$$B \sim \begin{pmatrix} 1 & 4 & -3 \\ 0 & 0 & 7 \\ 0 & 11 & -8 \end{pmatrix}$$

Obsérvese que la segunda entrada diagonal es 0. Siendo así, B no podrá llevarse a forma triangular sin operaciones de intercambio de filas. Dicho de otro modo, B no es factorizable LU.

4.69. Encontrar la factorización LU de $A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & -8 & 5 \\ 1 & 3 & 1 & 3 \\ 3 & 8 & -1 & 13 \end{pmatrix}$ por un método directo.

Comenzamos por construir las siguientes matrices L y U:

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

La parte del producto LU que determina la primera fila de A conduce a las cuatro ecuaciones

$$u_{11} = 1 \qquad u_{12} = 2 \qquad u_{13} = -3 \qquad u_{14} = 4$$

y la que determina la primera columna a las ecuaciones

$$l_{21}u_{11} = 2, \qquad l_{31}u_{11} = 1, \qquad l_{41}u_{11} = 3 \qquad \text{o} \qquad l_{21} = 2, \qquad l_{31} = 1, \qquad l_{41} = 3$$

Llegando a este punto, las matrices L y U presentan la forma

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & l_{32} & 1 & 0 \\ 3 & l_{42} & l_{43} & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

La parte del producto LU que determina el resto de las entradas en la segunda fila de A proporciona las ecuaciones

$$4 + u_{22} = 3 \qquad -6 + u_{23} = -8 \qquad 8 + u_{24} = 5$$

o

$$u_{22} = -1 \qquad u_{23} = -2 \qquad u_{24} = -3$$

y la que determina el resto de las entradas de la segunda columna nos lleva a

$$2 + l_{32}u_{22} = 3, \qquad 6 + l_{42}u_{22} = 8 \qquad \text{o} \qquad l_{32} = -1, \qquad l_{42} = -2$$

De este modo, L y U tienen ahora la forma

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 3 & -2 & l_{43} & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

Continuando con la tercera fila, tercera columna y cuarta fila de A obtenemos

$$u_{33} = 2, \qquad u_{34} = -1, \qquad \text{después} \qquad l_{43} = 2, \qquad \text{y, por último,} \qquad u_{44} = 3$$

Así

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 3 & -2 & 2 & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

4.70. Hallar la factorización LDU de la matriz A del Problema 4.69.

Aquí U debe tener unos en la diagonal y D ser una matriz diagonal. Así, usando la factorización LU precedente, sacamos como factores las entradas diagonales de aquella U llegando a

$$D = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 2 & \\ & & & 3 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 2 & -3 & 4 \\ & 1 & 2 & 3 \\ & & 1 & -2 \\ & & & 1 \end{pmatrix}$$

La matriz L coincide con la del Problema 4.69.

4.71. Se da la factorización $A = LU$, donde $L = (l_{ij})$ y $U = (u_{ij})$. Considérese el sistema $AX = B$. Se pide determinar: *a*) el algoritmo para encontrar $L^{-1}B$, *b*) el algoritmo que resuelve $UX = B$ vía sustitución hacia atrás.

a) La entrada l_{ij} de la matriz L corresponde a la operación elemental entre filas $-l_{ij}R_i + R_j \rightarrow R_j$. Por ello, el algoritmo que transforma B en B' es:

***Algoritmo** P4.88A: Evaluación de $L^{-1}B$.*

Paso 1. Repetir para $j = 1$ a $n - 1$:

Paso 2. Repetir para $i = j + 1$ a n:

$$b_j := -l_{ij}b_i + b_j$$

[Fin del bucle interno Paso 2.]

[Fin del bucle externo Paso 1.]

Paso 3. Salir.

[La complejidad de este algoritmo es $C(n) \approx n^2/2$.]

b) El algoritmo de sustitución hacia atrás es el que se expone a continuación:

***Algoritmo** P4.88B: Sustitución hacia atrás para el sistema $UX = B$.*

Paso 1. $x_n = b_n/u_{nn}$.

Paso 2. Repetir para $j = n - 1, n - 2, \ldots, 1$

$$x_j = (b_j - u_{j,j+1}x_{j+1} - \cdots - u_{jn}x_n)/u_{jj}$$

Paso 3. Salir.

[También aquí la complejidad es $C(n) \approx n^2/2$.]

4.72. Encontrar la factorización LU de la matriz $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & -10 & 2 \end{pmatrix}$.

Reducimos A a forma triangular mediante las operaciones

1. $-2R_1 + R_2 \rightarrow R_2$, 2. $3R_1 + R_3 \rightarrow R_3$, 3. $-4R_2 + R_3 \rightarrow R_3$

$$A \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & -4 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Por tanto, $$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Las entradas 2, -3 y 4 de L son los opuestos de los multiplicadores en las operaciones entre filas precedentes.

4.73. Resolver el sistema $AX = B$ para B_1, B_2 y B_3, donde A es la matriz del Problema 4.72 y $B_1 = (1, 1, 1)$, $B_2 = B_1 + X_1$, $B_3 = B_2 + X_2$ (aquí X_j es la solución cuando $B = B_j$).

a) Calculamos $L^{-1}B_1$ o, equivalentemente, efectuamos las operaciones entre filas (1), (2) y (3) sobre B_1, lo que nos conduce a

$$B_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{(1)\text{ y }(2)} \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}$$

Resolvemos $UX = B$ para $B = (1, -1, 8)^T$ por sustitución hacia atrás obteniendo $X_1 = (-25, 9, 8)^T$.

b) Hallamos $B_2 = B_1 + X_1 = (1, 1, 1) + (-25, 9, 8) = (-24, 10, 9)$. Efectuamos las operaciones (1), (2) y (3) sobre B_2 llegando a $(-24, 58, -63)^T$ y luego a $B = (-24, 58, -295)^T$.
Resolvemos $UX = B$ por sustitución hacia atrás obteniendo $X_2 = (943, -353, -295)$.

c) Calculamos $B_3 = B_2 + X_2 = (-24, 10, 9) + (943, -353, -295) = (919, -343, -286)$. Efectuamos las operaciones (1), (2) y (3) sobre B_3 llegando a $(919, -2187, 2671)^T$ y después a $B = (919, -2181, 11.395)^T$.
Resolvemos $UX = B$ por sustitución hacia atrás obteniendo $X_3 = (-37.628, 13.576, 11.395)$.

PROBLEMAS SUPLEMENTARIOS

ALGEBRA DE MATRICES

4.74. Sea $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Calcular A^n.

4.75. Supóngase que la matriz 2×2 B conmuta con cualquier matriz 2×2 A. Probar que $B = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ para algún escalar k, es decir, B es una matriz escalar.

4.76. Sea $A = \begin{pmatrix} 5 & 2 \\ 0 & k \end{pmatrix}$. Encontrar todos los números k para los que A es una raíz del polinomio

a) $f(x) = x^2 - 7x + 10$, *b*) $g(x) = x^2 - 25$, *c*) $h(x) = x^2 - 4$

4.77. Sea $B = \begin{pmatrix} 1 & 0 \\ 26 & 27 \end{pmatrix}$. Hallar una matriz A tal que $A^3 = B$.

4.78. Sean $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Calcular: *a*) A^n para todos los enteros positivos n, *b*) B^n para todos los enteros positivos n.

4.79. Imponer sobre las matrices A y B las condiciones necesarias para que $A^2 - B^2 = (A + B)(A - B)$.

MATRICES INVERTIBLES. INVERSAS. MATRICES ELEMENTALES

4.80. Hallar la inversa de cada una de las matrices:

$$a)\begin{pmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{pmatrix},\quad b)\begin{pmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{pmatrix},\quad c)\begin{pmatrix} 1 & -2 & 0 \\ 2 & -3 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

4.81. Hallar la inversa de cada una de las matrices:

$$a)\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix},\quad b)\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 3 & 1 & -2 \\ 1 & 4 & -2 & 4 \end{pmatrix}$$

4.82. Expresar cada una de las siguientes matrices como producto de matrices elementales: *a*) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, *b*) $\begin{pmatrix} 3 & -6 \\ -2 & 4 \end{pmatrix}$.

4.83. Expresar $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 3 & 8 & 7 \end{pmatrix}$ como producto de matrices elementales.

4.84. Supóngase que A es invertible. Demostrar que si $AB = AC$, necesariamente $B = C$. Dar un ejemplo de una matriz no nula A tal que $AB = AC$ pero $B \neq C$.

4.85. Si A es invertible, demostrar que kA es invertible para $k \neq 0$, con inversa $k^{-1}A^{-1}$.

4.86. Supóngase que A y B son invertibles y que $A + B \neq 0$. Probar, con un ejemplo, que $A + B$ no es necesariamente invertible.

TIPOS ESPECIALES DE MATRICES CUADRADAS

4.87. Utilizando sólo los elementos 0 y 1, hallar todas las matrices triangulares superiores 3×3 no singulares.

4.88. Utilizando sólo los elementos 0 y 1, determinar el número de: *a*) matrices diagonales 4×4, *b*) matrices triangulares superiores 4×4, *c*) matrices triangulares superiores 4×4 no singulares. Generalizar el resultado a matrices $n \times n$.

4.89. Encontrar todas las matrices reales A tales que $A^2 = B$, donde: *a*) $B = \begin{pmatrix} 4 & 21 \\ 0 & 25 \end{pmatrix}$, *b*) $B = \begin{pmatrix} 1 & 4 \\ 0 & -9 \end{pmatrix}$.

4.90. Sea $B = \begin{pmatrix} 1 & 8 & 5 \\ 0 & 9 & 5 \\ 0 & 0 & 4 \end{pmatrix}$. Hallar una matriz A con entradas diagonales positivas tal que $A^2 = B$.

4.91. Supóngase $AB = C$, siendo A y C triangulares superiores.

a) Demostrar, con un ejemplo, que B no es necesariamente triangular superior, ni siquiera cuando A y C son matrices no nulas.
b) Demostrar que B es triangular superior cuando A es invertible.

4.92. Probar que AB no es necesariamente simétrica incluso en el caso de que A y B lo sean.

4.93. Sean A y B matrices simétricas. Demostrar que AB es simétrica si y sólo si A y B conmutan.

4.94. Supóngase que A es una matriz simétrica. Probar que: *a*) A^2 y, en general, A^n son simétricas; *b*) $f(A)$ es simétrica para cualquier polinomio $f(x)$; *c*) $P^T AP$ es simétrica.

4.95. Encontrar una matriz ortogonal 2×2 P cuya primera fila sea *a*) $(2/\sqrt{29}, 5/\sqrt{29})$, *b*) un múltiplo de $(3, 4)$.

4.96. Hallar una matriz ortogonal 3×3 P cuyas dos primeras filas sean múltiplos de *a*) $(1, 2, 3)$ y $(0, -2, 3)$, respectivamente; *b*) $(1, 3, 1)$ y $(2, 0, -1)$, respectivamente.

4.97. Supóngase que A y B son ortogonales. Probar que A^T, A^{-1} y AB también lo son.

4.98. ¿Cuáles de las siguientes matrices son normales?

$$A = \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & 1 \\ -3 & -1 & 2 \end{pmatrix}$$

4.99. Supóngase que A es una matriz normal. Demostrar que: *a*) A^T, *b*) A^2 y, en general, A^n, *c*) $B = kI + A$ también lo son.

4.100. Una matriz E es *idempotente* si $E^2 = E$. Probar que $E = \begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix}$ es idempotente.

4.101. Demostrar que si $AB = A$ y $BA = B$, entonces A y B son idempotentes.

4.102. Una matriz A es *nilpotente de clase p* si $A^p = 0$ pero $A^{p-1} \neq 0$. Probar que $A = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}$ es nilpotente de clase 3.

4.103. Supóngase que A es nilpotente de clase p. Probar que $A^q = 0$ para $q > p$ pero $A^q \neq 0$ para $q < p$.

4.104. Una matriz cuadrada es *tridiagonal* si las entradas no nulas aparecen únicamente en la diagonal situada directamente sobre la diagonal principal (en la superdiagonal), o directamente bajo la diagonal principal (en la subdiagonal). Exhibir las matrices tridiagonales genéricas de órdenes 4 y 5.

4.105. Probar que el producto de matrices tridiagonales no es necesariamente tridiagonal.

MATRICES COMPLEJAS

4.106. Encontrar tres números reales x, y y z tales que A sea hermítica, siendo

$$a)\quad A = \begin{pmatrix} x + yi & 3 \\ 3 + zi & 0 \end{pmatrix}, \qquad b)\quad A = \begin{pmatrix} 3 & x + 2i & yi \\ 3 - 2i & 0 & 1 + zi \\ yi & 1 - xi & -1 \end{pmatrix}$$

4.107. Supóngase que A es una matriz compleja arbitraria. Demostrar que AA^H y $A^H A$ son ambas hermíticas.

4.108. Supóngase que A es cualquier matriz compleja. Demostrar que $A + A^H$ es hermítica y que $A - A^H$ es antihermítica.

4.109. ¿Cuáles de las siguientes matrices son unitarias?

$$A = \begin{pmatrix} i/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -i/2 \end{pmatrix}, \quad B = \frac{1}{2}\begin{pmatrix} 1 + i & 1 - i \\ 1 - i & 1 + i \end{pmatrix}, \quad C = \frac{1}{2}\begin{pmatrix} 1 & -i & -1 + i \\ i & 1 & 1 + i \\ 1 + i & -1 + i & 0 \end{pmatrix}$$

4.110. Supóngase que A y B son matrices unitarias. Probar que: *a*) A^H es unitaria, *b*) A^{-1} es unitaria, *c*) AB es unitaria.

4.111. Determinar cuáles de las siguientes matrices son normales: $A = \begin{pmatrix} 3 + 4i & 1 \\ i & 2 + 3i \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 - i & i \end{pmatrix}$.

4.112. Supóngase que A es una matriz normal y U una unitaria. Probar que $B = U^H AU$ es también normal.

4.113. Recuérdense las siguientes operaciones elementales entre filas:

$$[E_1]\ R_i \leftrightarrow R_j, \quad [E_2]\ kR_i \to R_i, \quad k \neq 0, \quad [E_3]\ kR_j + R_i \to R_i$$

Para matrices complejas, las respectivas operaciones entre columnas hermíticas correspondientes son:

$$[G_1]\ C_i \leftrightarrow C_j, \quad [G_2]\ \bar{k}C_i \to C_i, \quad k \neq 0, \quad [G_3]\ \bar{k}C_j + C_i \to C_i$$

Demostrar que la matriz elemental correspondiente a $[G_i]$ es la traspuesta conjugada de la matriz elemental asociada a $[E_i]$.

MATRICES CUADRADAS POR BLOQUES

4.114. Usando líneas verticales, completar la partición de cada una de las siguientes matrices para que sea una matriz cuadrada por bloques:

$$A = \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ \hdashline 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 7 & 6 & 5 \\ \hdashline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right), \qquad B = \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hdashline 9 & 8 & 7 & 6 & 5 \\ \hdashline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right)$$

4.115. Partir cada una de las siguientes matrices para convertirla en una matriz diagonal por bloques, con tantos bloques diagonales como sea posible:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

4.116. Calcular M^2 y M^3 para cada matriz M:

$$a) \quad M = \left(\begin{array}{c:cc:c} 2 & 0 & 0 & 0 \\ \hdashline 0 & 1 & 4 & 0 \\ 0 & 2 & 1 & 0 \\ \hdashline 0 & 0 & 0 & 3 \end{array}\right), \qquad b) \quad M = \left(\begin{array}{cc:cc} 1 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 2 \\ 0 & 0 & 4 & 5 \end{array}\right).$$

4.117. Sean $M = \operatorname{diag}(A_1, \ldots, A_k)$ y $N = \operatorname{diag}(B_1, \ldots, B_k)$ matrices diagonales por bloques, donde cada par de bloques A_i, B_i tienen el mismo tamaño. Probar que MN es diagonal por bloques y que

$$MN = \operatorname{diag}(A_1 B_1, A_2 B_2, \ldots, A_k B_k)$$

MATRICES REALES SIMETRICAS Y FORMAS CUADRATICAS

4.118. Sea $A = \begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -5 & -1 \\ -2 & -5 & 6 & 9 \\ -3 & -1 & 9 & 11 \end{pmatrix}$. Hallar una matriz no singular P tal que $B = P^T A P$ sea diagonal. Asimismo, determinar B y $\operatorname{sig} A$.

4.119. Para cada forma cuadrática $q(x, y, z)$, encontrar una sustitución lineal no singular que exprese las variables x, y, z en términos de variables r, s, t tales que $q(r, s, t)$ sea diagonal.

a) $q(x, y, z) = x^2 + 6xy + 8y^2 - 4xz + 2yz - 9z^2$.

b) $q(x, y, z) = 2x^2 - 3y^2 + 8xz + 12yz + 25z^2$.

4.120. Encontrar aquellos valores de k para los que la forma cuadrática dada es definida positiva:

a) $q(x, y) = 2x^2 - 5xy + ky^2$.

b) $q(x, y) = 3x^2 - kxy + 12y^2$.

c) $q(x, y, z) = x^2 + 2xy + 2y^2 + 2xz + 6yz + kz^2$.

4.121. Dar un ejemplo de una forma cuadrática $q(x, y)$ tal que $q(u) = 0$ y $q(v) = 0$ pero $q(u + v) \neq 0$.

4.122. Demostrar que cualquier matriz real simétrica A es congruente a una matriz diagonal en la que cada entrada diagonal es 1, -1 ó 0.

4.123. Demostrar que la congruencia de matrices es una relación de equivalencia.

SIMILARIDAD DE MATRICES

4.124. Considérese el espacio $\mathbf{R}^3$ con los ejes x, y, z usuales. La matriz no singular $P = \begin{pmatrix} 1 & -2 & -2 \\ 2 & -3 & -6 \\ 1 & 1 & -7 \end{pmatrix}$ determina un nuevo sistema de coordenadas para $\mathbf{R}^3$, digamos con ejes r, s, t. Hallar:

a) Las coordenadas del punto $Q(1, 1, 1)$ en el nuevo sistema.

b) $f(r, s, t)$ cuando $f(x, y, z) = (x + y, y + 2z, x - z)$.

c) $g(r, s, t)$ cuando $g(x, y, z) = (x + y - z, x - 3z, 2x + y)$.

4.125. Demostrar que la similaridad de matrices es una relación de equivalencia.

FACTORIZACION *LU*

4.126. Hallar las factorizaciones LU y LDU de cada matriz:

a) $A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 5 & 1 \\ 3 & 4 & 2 \end{pmatrix}$, b) $B = \begin{pmatrix} 2 & 3 & 6 \\ 4 & 7 & 9 \\ 3 & 5 & 4 \end{pmatrix}$.

4.127. Sea $A = \begin{pmatrix} 1 & -1 & -1 \\ 3 & -4 & -2 \\ 2 & -3 & -2 \end{pmatrix}$.

a) Encontrar la factorización LU de A.

b) Denótese por X_k la solución de $AX = B_k$. Calcular X_1, X_2, X_3, X_4 cuando $B_1 = (1, 1, 1)^T$ y $B_{k+1} = B_k + X_k$ para $k > 0$.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

4.74. $\begin{pmatrix} 1 & 2^n \\ 0 & 1 \end{pmatrix}$

4.76. *a*) $k = 2$, *b*) $k = -5$, *c*) ninguno.

4.77. $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$

4.78. *a*) $A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $A^k = 0$ para $k > 3$; *b*) $B^n = \begin{pmatrix} 1 & n & n(n-1)/2 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$

4.79. $AB = BA$.

4.80. *a*) $\begin{pmatrix} \frac{29}{2} & -\frac{17}{2} & \frac{7}{2} \\ -\frac{5}{2} & \frac{3}{2} & -\frac{1}{2} \\ 3 & -2 & 1 \end{pmatrix}$, *b*) $\begin{pmatrix} 8 & -3 & -1 \\ -5 & 2 & 1 \\ 10 & -4 & -1 \end{pmatrix}$, *c*) $\begin{pmatrix} -8 & 5 & -1 \\ -\frac{9}{2} & \frac{5}{2} & -\frac{1}{2} \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$

4.81. *a*) $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, *b*) $\begin{pmatrix} -10 & -20 & 4 & 7 \\ [illegible] & 6 & -1 & -2 \\ [illegible] & 8 & -2 & -3 \\ 2 & 3 & -1 & -1 \end{pmatrix}$

4.82. *a*) $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$ o $\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. *b*) No hay tal producto: la matriz no tiene inversa.

4.83. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

4.84. $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$

4.86. $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}$

4.87. Todas las entradas diagonales deben ser 1 para que sean no singulares. Hay ocho elecciones posibles para las entradas sobre la diagonal:

$$\begin{pmatrix}0 & 0\\ * & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ * & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ * & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ * & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ * & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ * & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ * & 0\end{pmatrix}, \begin{pmatrix}1 & 1\\ * & 1\end{pmatrix}$$

4.88. *a)* $2^4[2^n]$, *b)* $2^{10}[2^{n(n+1)/2}]$, *c)* $2^6[2^{n(n-1)/2}]$.

4.89. *a)* $\begin{pmatrix}2 & 4\\ 0 & 5\end{pmatrix}, \begin{pmatrix}2 & -7\\ 0 & -5\end{pmatrix}, \begin{pmatrix}-2 & 7\\ 0 & 5\end{pmatrix}, \begin{pmatrix}-2 & -4\\ 0 & -5\end{pmatrix}$, *b)* ninguna.

4.90. $\begin{pmatrix}1 & 2 & 1\\ & 3 & 1\\ & & 2\end{pmatrix}$

4.91. *a)* $A = \begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}, B = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}, C = \begin{pmatrix}4 & 6\\ 0 & 0\end{pmatrix}$

4.92. $\begin{pmatrix}1 & 2\\ 2 & 2\end{pmatrix}\begin{pmatrix}3 & 3\\ 3 & 1\end{pmatrix} = \begin{pmatrix}9 & 5\\ 12 & 8\end{pmatrix}$

4.95. *a)* $\begin{pmatrix}2/\sqrt{29} & 5/\sqrt{29}\\ -5/\sqrt{29} & 2/\sqrt{29}\end{pmatrix}$, *b)* $\begin{pmatrix}\frac{3}{5} & \frac{4}{5}\\ -\frac{4}{5} & \frac{3}{5}\end{pmatrix}$

4.96. *a)* $\begin{pmatrix}1/\sqrt{14} & 2/\sqrt{14} & 3/\sqrt{14}\\ 0 & -2/\sqrt{13} & 3/\sqrt{13}\\ 12/\sqrt{157} & -3/\sqrt{157} & -2/\sqrt{157}\end{pmatrix}$, *b)* $\begin{pmatrix}1/\sqrt{11} & 3/\sqrt{11} & 1/\sqrt{11}\\ 1/\sqrt{2} & 0 & -1/\sqrt{2}\\ 3/\sqrt{22} & -2/\sqrt{22} & 3/\sqrt{22}\end{pmatrix}$

4.98. A, C.

4.104. $\begin{pmatrix}a_{11} & a_{21} & & \\ a_{21} & a_{22} & a_{32} & \\ & a_{32} & a_{33} & a_{34}\\ & & a_{43} & a_{44}\end{pmatrix}$ $\begin{pmatrix}b_{11} & b_{21} & & & \\ b_{21} & b_{22} & b_{23} & & \\ & b_{32} & b_{33} & b_{34} & \\ & & b_{43} & b_{44} & b_{45}\\ & & & b_{54} & b_{55}\end{pmatrix}$

4.105. $\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 1\\ 0 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 1\\ 0 & 1 & 1\end{pmatrix} = \begin{pmatrix}2 & 2 & 1\\ 2 & 3 & 2\\ 1 & 2 & 2\end{pmatrix}$

4.106. *a)* $x = a$ (parámetro), $y = 0$, $z = 0$; *b)* $x = 3$, $y = 0$, $z = 3$.

4.109. A, B, C.

4.111. A.

4.114. $A = \left(\begin{array}{c|cc|cc} 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right), \quad B = \left(\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 9 & 8 & 7 & 6 & 5 \\ \hline 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \end{array}\right)$

4.115. $A = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 2 \\ 0 & 0 & 3 \end{array}\right), \quad B = \left(\begin{array}{cc|ccc} 1 & 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 6 \end{array}\right), \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$

(C ya es una matriz diagonal por bloques; no es posible ninguna partición ulterior de C.)

4.116. *a)* $M^2 = \begin{pmatrix} 4 & & & \\ & 9 & 8 & \\ & 4 & 9 & \\ & & & 9 \end{pmatrix}, \quad M^3 = \begin{pmatrix} 8 & & & \\ & 25 & 44 & \\ & 22 & 25 & \\ & & & 27 \end{pmatrix}$

b) $M^2 = \begin{pmatrix} 3 & 4 & & \\ 8 & 11 & & \\ & & 9 & 12 \\ & & 24 & 33 \end{pmatrix}, \quad M^3 = \begin{pmatrix} 11 & 15 & & \\ 30 & 41 & & \\ & & 57 & 78 \\ & & 156 & 213 \end{pmatrix}$

4.118. $P = \begin{pmatrix} 1 & -1 & -1 & 26 \\ 0 & 1 & 3 & 13 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -7 & \\ & & & 469 \end{pmatrix}, \quad \text{sig } A = 2$

4.119. *a)* $x = r - 3s + 19t$, $y = s + 7t$, $z = t$,
$q(r, s, t) = r^2 - s^2 + 36t^2$, rango $q = 3$, sig $q = 1$.

b) $x = r - 2t$, $y = s + 2t$, $z = t$,
$q(r, s, t) = 2r^2 - 3s^2 + 29t^2$, rango $q = 3$, sig $q = 1$.

c) $x = r - 2s + 18t$, $y = s - 7t$, $z = t$,
$q(r, s, t) = r^2 + s^2 - 62t^2$, rango $q = 3$, sig $q = 1$.

d) $x = r - s - t$, $y = s - t$, $z = t$,
$q(x, y, z) = r^2 + 2s^2$, rango $q = 2$, sig $q = 2$.

4.120. *a)* $k > \frac{25}{8}$; *b)* $k < -12$ ó $k > 12$; *c)* $k > 5$.

4.121. $q(x, y) = x^2 - y^2$, $u = (1, 1)$, $v = (1, -1)$.

4.122. Supongamos que A se ha diagonalizado a $P^T AP = \text{diag}\,(a_i)$. Defínase $Q = \text{diag}\,(b_i)$ por $b_i = \begin{cases} 1/\sqrt{|a_i|} & \text{si } a_i \neq 0 \\ 1 & \text{si } a_i = 0 \end{cases}$. Entonces $B = Q^T P^T APQ = (PQ)^T A(PQ)$ tiene la forma requerida.

4.124. *a)* $Q(17, 5, 3)$, (*b*) $f(r, s, t) = (17r - 61s + 134t, 4r - 41s + 46t, 3r - 25s + 25t)$,
c) $g(r, s, t) = (61r + s - 330t, 16r + 3s - 91t, 9r - 4s - 4t)$.

4.126. *a)* $A = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ 3 & 5 & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & -1 & \\ & & -10 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 \\ & 1 & -3 \\ & & 1 \end{pmatrix}$

b) $B = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ \frac{3}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & & \\ & 1 & \\ & & -\frac{7}{2} \end{pmatrix} \begin{pmatrix} 1 & \frac{3}{2} & 3 \\ & 1 & -3 \\ & & 1 \end{pmatrix}$

4.127. *a)* $A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$

b) $X_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $B_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$, $X_2 = \begin{pmatrix} 6 \\ 4 \\ 0 \end{pmatrix}$, $B_3 = \begin{pmatrix} 8 \\ 6 \\ 0 \end{pmatrix}$, $X_3 = \begin{pmatrix} 22 \\ 16 \\ -2 \end{pmatrix}$, $B_4 = \begin{pmatrix} 30 \\ 22 \\ -2 \end{pmatrix}$, $X_4 = \begin{pmatrix} 86 \\ 62 \\ -6 \end{pmatrix}$

CAPITULO 5

Espacios vectoriales

5.1. INTRODUCCION

Este capítulo introduce la estructura algebraica subyacente al álgebra lineal, la de espacio vectorial de dimensión finita. La definición de espacio vectorial involucra un cuerpo arbitrario (véase el Apéndice) cuyos elementos se denominan *escalares*. Se utilizará la siguiente notación (a menos que se establezca o implique otra cosa):

K	el cuerpo de escalares
a, b, c o k	los elementos de K
V	el espacio vectorial dado
u, v, w	los elementos de V

No se perderá nada esencial si el lector supone que K es el cuerpo real **R** o el complejo **C**.

El capítulo no abarca conceptos como longitud y ortogonalidad, puesto que no se consideran parte de la estructura fundamental de un espacio vectorial. Se incluirán como estructura adicional en el Capítulo 6.

5.2. ESPACIOS VECTORIALES

A continuación se define la noción de espacio vectorial o lineal.

Definición: Sean K un cuerpo dado y V un conjunto no vacío, con reglas de suma y producto por un escalar que asignan a cada par $u, v \in V$ una *suma* $u + v \in V$ y a cada par $u \in V$, $k \in K$ un *producto* $ku \in V$. V recibe el nombre de *espacio vectorial sobre* K (y los elementos de V se llaman *vectores*) si se satisfacen los siguientes axiomas (véase el Problema 5.3).

[A_1] Para toda terna de vectores $u, v, w \in V$, $(u + v) + w = u + (v + w)$.

[A_2] Existe un vector en V, denotado por 0 y denominado el *vector cero*, tal que $u + 0 = u$ para todo vector $u \in V$.

[A_3] Para todo vector $u \in V$ existe un único vector en V, denotado por $-u$, tal que $u + (-u) = 0$.

[A_4] Para todo par de vectores $u, v \in V$, $u + v = v + u$.

[M_1] Para todo escalar $k \in K$ y todo par de vectores $u, v \in V$, $k(u + v) = ku + kv$.

[M_2] Para todo par de escalares $a, b \in K$ y todo vector $u \in V$, $(a + b)u = au + bu$.

[M_3] Para todo par de escalares $a, b \in K$ y todo vector $u \in V$, $(ab)u = a(bu)$.

[M_4] El escalar unidad $1 \in K$ cumple $1u = u$ para todo vector $u \in V$.

Los axiomas precedentes se desdoblan de forma natural en dos categorías. Los cuatro primeros atañen únicamente a la estructura aditiva de V y pueden resumirse diciendo que V es un *grupo conmutativo* (véase el Apéndice) bajo la suma. De ello se deriva que cualquier suma de vectores de la forma

$$v_1 + v_2 + \cdots + v_m$$

no requiere paréntesis y no depende del orden de los sumandos, que el vector cero, 0, es único, que el *opuesto* $-u$ de u es único y que se verifica la *ley de cancelación*; esto es, para tres vectores cualesquiera $u, v, w \in V$.

$$u + w = v + w \qquad \text{implica} \qquad u = v$$

Asimismo, la *resta* se define según

$$u - v = u + (-v)$$

Por otra parte, los cuatro axiomas restantes se refieren a la «acción» del cuerpo K sobre V. Obsérvese que la rotulación de los axiomas refleja este desdoblamiento. Empleando estos axiomas adicionales probaremos (Problema 5.1) las siguientes propiedades elementales de un espacio vectorial.

Teorema 5.1: Sea V un espacio vectorial sobre un cuerpo K.

i) Para todo escalar $k \in K$ y $0 \in V$, $k0 = 0$.

ii) Para $0 \in K$ y todo vector $u \in V$, $0u = 0$.

iii) Si $ku = 0$, donde $k \in K$ y $u \in V$, entonces $k = 0$ o $u = 0$.

iv) Para todo $k \in K$ y todo $u \in V$, $(-k)u = k(-u) = -ku$.

5.3. EJEMPLOS DE ESPACIOS VECTORIALES

Esta sección enumera una serie de ejemplos importantes de espacios vectoriales que se utilizarán a lo largo de todo el texto.

ESPACIO K^n

Sea K un cuerpo arbitrario. La notación K^n se usa frecuentemente para designar el conjunto de todas las n-plas de elementos de K. Aquí K^n se ve como un espacio vectorial sobre K, en el que la suma vectorial y el producto por un escalar se definen según

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

y

$$k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$$

El vector cero en K^n es la n-pla de ceros

$$0 = (0, 0, \ldots, 0)$$

y el opuesto de un vector se define por

$$-(a_1, a_2, \ldots, a_n) = (-a_1, -a_2, \ldots, -a_n)$$

La demostración de que K^n es un espacio vectorial es idéntica a la del Teorema 2.1, que ahora puede considerarse como la afirmación de que $\mathbf{R}^n$, con las operaciones allí definidas, es un espacio vectorial sobre $\mathbf{R}$.

ESPACIO DE MATRICES $\mathbf{M}_{m,n}$

La notación $\mathbf{M}_{m,n}$, o simplemente $\mathbf{M}$, se utilizará para designar el conjunto de todas las matrices $m \times n$ sobre un cuerpo arbitrario K. $\mathbf{M}_{m,n}$ es un espacio vectorial sobre K con respecto a las operaciones usuales de suma matricial y producto por un escalar. (Véase el Teorema 3.1.)

ESPACIO DE POLINOMIOS $\mathbf{P}(t)$

Denotemos por $\mathbf{P}(t)$ el conjunto de todos los polinomios

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$$

con coeficientes a_i en algún cuerpo K. (Véase el Apéndice.) $\mathbf{P}(t)$ es un espacio vectorial sobre K con respecto a las operaciones usuales de suma de polinomios y producto de un polinomio por una constante.

ESPACIO DE FUNCIONES $F(X)$

Sean X un conjunto no vacío y K un cuerpo arbitrario. Consideremos el conjunto $F(X)$ de todas las funciones de X en K. [Nótese que $F(X)$ es no vacío por serlo X.] La suma de dos funciones $f, g \in F(X)$ es la función $f + g \in F(X)$ definida por

$$(f + g)(x) = f(x) + g(x) \qquad \forall x \in X$$

y el producto de un escalar $k \in K$ por una función $f \in F(X)$ es la función $kf \in F(X)$ definida por

$$(kf)(x) = kf(x) \qquad \forall x \in X$$

(El símbolo $\forall$ significa «para todo».) $F(X)$, con las operaciones anteriores, es un espacio vectorial sobre K (Problema 5.5).

El vector cero en $F(X)$ es la función cero, $\mathbf{0}$, que aplica cada $x \in X$ en $0 \in K$, es decir,

$$\mathbf{0}(x) = 0 \qquad \forall x \in X$$

Asimismo, para cualquier función $f \in F(X)$, la función $-f$ definida por

$$(-f)(x) = -f(x) \qquad \forall x \in X$$

es la opuesta de la función f.

CUERPOS Y SUBCUERPOS

Supongamos que E es un cuerpo que contiene un subcuerpo K. E puede verse como un espacio vectorial sobre K como sigue. Tomemos como suma vectorial la suma usual en E, y como producto kv del escalar $k \in K$ por el vector $v \in E$ el producto de k y v como elementos del cuerpo E. En tal caso, E es un espacio vectorial sobre K, esto es, E y K satisfacen los ocho axiomas de espacio vectorial precedentes.

5.4. SUBESPACIOS

Sea W un subconjunto de un espacio vectorial V sobre un cuerpo K. W se denomina un *subespacio* de V si es a su vez un espacio vectorial sobre K con respecto a las operaciones de V, suma vectorial y producto por un escalar. Un criterio simple para identificar subespacios es el siguiente (véase el Problema 5.4 para su demostración).

Teorema 5.2: Supongamos que W es un subconjunto de un espacio vectorial V. Entonces W es un subespacio de V si y sólo si se cumple:

i) $0 \in W$.

ii) W es *cerrado bajo la suma de vectores*, es decir:
Para todo par de vectores $u, v \in W$, la suma $u + v \in W$.

iii) W es *cerrado bajo el producto por un escalar*, esto es:
Para todo $u \in W$ y para todo $k \in K$ el múltiplo $ku \in W$.

Las condiciones ii) y iii) pueden combinarse en una sola, como se hace, más abajo, en ii) (véase la demostración en el Problema 5.5).

Corolario 5.3: W es un subespacio de V si y sólo si:

i) $0 \in W$.

ii) $au + bv \in W$ para todos los $u, v \in W$ y $a, b \in K$.

EJEMPLO 5.1

a) Sea V cualquier espacio vectorial. Entonces, tanto el conjunto $\{\mathbf{0}\}$, que consiste en el vector cero solo, como el espacio V entero son subespacios de V.

b) Sea W el plano xy en $\mathbf{R}^3$ consistente en aquellos vectores cuya tercera componente es 0; o, en otras palabras, $W = \{(a, b, 0) : a, b \in \mathbf{R}\}$. Nótese que $0 = (0, 0, 0) \in W$ ya que la tercera componente de 0 es 0. Para todo par de vectores $u = (a, b, 0)$ y $v = (c, d, 0)$ en W y todo escalar $k \in \mathbf{R}$ tenemos que

$$u + v = (a + c, b + d, 0) \qquad \text{y} \qquad ku = (ka, kb, 0)$$

pertenecen a W, luego W es un subespacio de $\mathbf{R}^3$.

c) Sea $V = \mathbf{M}_{n,n}$ el espacio de las matrices $n \times n$. El subconjunto W_1 de las matrices triangulares (superiores) y el W_2 de las matrices simétricas son subespacios de V puesto que son no vacíos y cerrados bajo la suma de matrices y el producto por un escalar.

d) Recordemos que $\mathbf{P}(t)$ denota el espacio vectorial de los polinomios. Denotemos por $\mathbf{P}_n(t)$ el subconjunto de $\mathbf{P}(t)$ que consiste en todos los polinomios de grado $\leq n$, para un n fijo. En ese caso, $\mathbf{P}_n(t)$ es un subespacio de $\mathbf{P}(t)$. El espacio vectorial $\mathbf{P}_n(t)$ aparecerá muy a menudo en nuestros ejemplos.

EJEMPLO 5.2. Sean U y W subespacios de un espacio vectorial V. Probemos que la intersección $U \cap W$ es también subespacio de V. Claramente, $0 \in U$ y $0 \in W$, porque U y W son subespacios, de donde $0 \in U \cap W$. Supongamos ahora que $u, v \in U \cap W$. Entonces $u, v \in U$ y $u, v \in E$ y, dado que U y W son subespacios,

$$u + v, \qquad ku \in U \qquad \text{y} \qquad u + v, \qquad ku \in W$$

para cualquier escalar k. Así $u + v$, $ku \in U \cap W$ y por consiguiente $U \cap W$ es subespacio de V.

El resultado del ejemplo precedente se generaliza como sigue.

Teorema 5.4: La intersección de cualquier número de subespacios de un espacio vectorial V es un subespacio de V.

Recuérdese que toda solución de un sistema de ecuaciones lineales con n incógnitas $AX = B$ puede verse como un punto en K^n y por tanto el conjunto solución de tal sistema es un subconjunto de K^n. Supongamos que el sistema es homogéneo, es decir, supongamos que el sistema tiene la forma $AX = 0$. Denotemos por W su conjunto solución. Como $A0 = 0$, el vector cero $0 \in W$. Además, si u y v pertenecen a W, esto es, si u y v son soluciones de $AX = 0$, necesariamente $Au = 0$ y $Av = 0$. Por esta razón, para todo par de escalares a y b en K, tendremos

$$A(au + bv) = aAu + bAv = a0 + b0 = 0 + 0 = 0$$

De esta manera, $au + bv$ es también una solución de $AX = 0$ o, dicho de otro modo, $au + bv \in W$. En consecuencia, según el Corolario 5.3, hemos demostrado:

Teorema 5.5: El conjunto solución W de un sistema homogéneo con n incógnitas $AX = 0$ es un subespacio de K^n.

Hacemos énfasis en que el conjunto solución de un sistema inhomogéneo $AX = B$ no es subespacio de K^n. De hecho, el vector cero, 0, no pertenece a dicho conjunto solución.

5.5. COMBINACIONES LINEALES. ENVOLVENTES LINEALES

Sean V un espacio vectorial sobre un cuerpo K y $v_1, v_2, \ldots, v_m \in V$. Cualquier vector en V de la forma

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m$$

donde los $a_i \in K$, recibe el nombre de *combinación lineal* de $v_1, v_2, \ldots, v_m$. El conjunto de todas las combinaciones lineales semejantes, denotado por

$$\text{lin}\,(v_1, v_2, \ldots, v_m)$$

se denomina *envolvente lineal* de $v_1, v_2, \ldots, v_m$.

En general, para cualquier subconjunto S de V, lin $S = \{0\}$ si S es vacío y lin S consiste en todas las combinaciones lineales de vectores de S.

El siguiente teorema se demostrará en el Problema 5.16.

Teorema 5.6: Sea S un subconjunto de un espacio vectorial V.

i) Entonces lin S es un subespacio de V que contiene a S.

ii) Si W es un subespacio de V que contiene a S, necesariamente lin $S \subseteq W$.

Por otra parte, dado un espacio vectorial V, se dice que los vectores $u_1, u_2, \ldots, u_r$ *generan* o forman un *conjunto generador* de V si

$$V = \text{lin}\,(u_1, u_2, \ldots, u_r)$$

En otras palabras, $u_1, u_2, \ldots, u_r$ generan V si, para todo $v \in V$, existen escalares $a_1, a_2, \ldots, a_r$ tales que

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_r u_r$$

esto es, si v es una combinación lineal de $u_1, u_2, \ldots, u_r$.

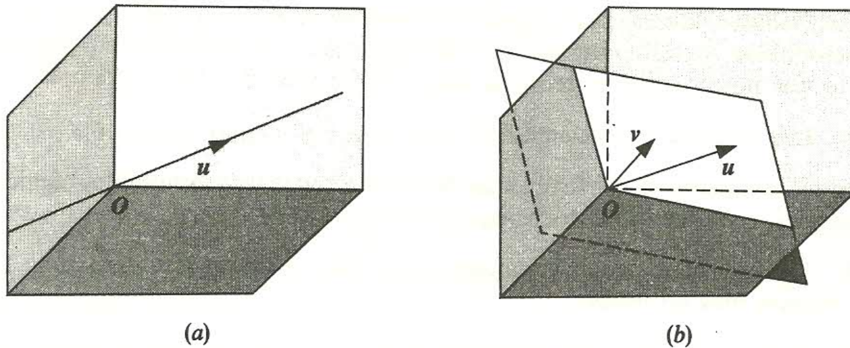
EJEMPLO 5.3

a) Consideremos el espacio vectorial $\mathbf{R}^3$. La envolvente lineal de cualquier vector no nulo $u \in \mathbf{R}^3$ consiste en todos los múltiplos escalares de u; geométricamente, lin u es la recta que pasa por el origen y por el extremo de u, tal y como se muestra en la Figura 5-1(*a*). Asimismo, para todo par de vectores $u, v \in \mathbf{R}^3$ que no sean uno múltiplo del otro, lin (u, v) es el plano que pasa por el origen y los extremos de u y v, como se muestra en la Figura 5-1(*b*).

b) Los vectores $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ y $e_3 = (0, 0, 1)$ generan el espacio vectorial $\mathbf{R}^3$. Concretamente, para todo vector $u = (a, b, c)$ en $\mathbf{R}^3$, tenemos

$$u = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = ae_1 + be_2 + ce_3$$

O sea, u es una combinación lineal de e_1, e_2, e_3.



(*a*) (*b*)

Figura 5-1.

c) Los polinomios $1, t, t^2, t^3, \ldots$ generan el espacio vectorial $\mathbf{P}(t)$ de todos los polinomios, es decir,

$$\mathbf{P}(t) = \text{lin}\,(1, t, t^2, t^3, \ldots)$$

Dicho de otro modo, todo polinomio es una combinación lineal de 1 y potencias de t. Similarmente, los polinomios $1, t, t^2, \ldots, t^n$ generan el espacio vectorial $\mathbf{P}_n(t)$ de los polinomios de grado $\leq n$.

ESPACIO FILA DE UNA MATRIZ

Sea A una matriz $m \times n$ sobre un cuerpo K arbitraria:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

Las filas de A,

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

pueden verse como vectores en K^n y por tanto generan un subespacio de K^n llamado el *espacio fila* de A y denotado por f-lin A. Esto es,

$$\text{f-lin } A = \text{lin } (R_1, R_2, \ldots, R_m)$$

Análogamente, las columnas de A pueden verse como vectores en K^m y por consiguiente generan un subespacio de K^m denominado el *espacio columna* de A y denotado por c-lin A. Alternativamente, c-lin A = f-lin A^T.

Supongamos ahora que efectuamos una operación elemental entre filas sobre A,

i) $R_i \leftrightarrow R_j$, ii) $kR_i \rightarrow R_i$, $k \neq 0$, o iii) $kR_j + R_i \rightarrow R_i$

y obtenemos una matriz B. En ese caso, cada fila de B es, claramente, bien una fila de A, o bien una combinación lineal de filas de A, por lo que el espacio fila de B está contenido en el de A.

Por otra parte, podemos efectuar la operación elemental entre filas inversa sobre B y obtener A, luego el espacio fila de A está contenido en el de B. De acuerdo con esto, A y B tienen el mismo espacio fila, lo que nos conduce al teorema enunciado a continuación.

Teorema 5.7: Las matrices equivalentes por filas tienen el mismo espacio fila.

En particular, demostraremos (Problemas 5.51 y 5.52, respectivamente) los siguientes resultados fundamentales, referentes a matrices equivalentes por filas.

Teorema 5.8: Dos matrices en forma canónica por filas tienen el mismo espacio fila si y sólo si tienen las mismas filas no nulas.

Teorema 5.9: Toda matriz es equivalente por filas a una única matriz en forma canónica por filas.

En el próximo ejemplo aplicamos los resultados anteriores.

EJEMPLO 5.4. Demostremos que el subespacio U de $\mathbf{R}^4$ generado por los vectores

$$u_1 = (1, 2, -1, 3) \qquad u_2 = (2, 4, 1, -2) \qquad \text{y} \qquad u_3 = (3, 6, 3, -7)$$

y el subespacio W de $\mathbf{R}^4$ generado por los vectores

$$v_1 = (1, 2, -4, 11) \qquad \text{y} \qquad v_2 = (2, 4, -5, 14)$$

son iguales, esto es, que $U = W$.

Método 1. Demostramos que cada u_i es combinación lineal de v_1 y v_2 y que cada v_i lo es de u_1, u_2 y u_3. Obsérvese que debemos probar la compatibilidad de seis sistemas de ecuaciones lineales.

Método 2. Construimos la matriz A, cuyas filas son los u_i, y la reducimos a forma canónica por filas:

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Ahora construimos la matriz B, cuyas filas son v_1 y v_2, y la reducimos a forma canónica por filas:

$$B = \begin{pmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{pmatrix}$$

Como las filas no nulas de las matrices reducidas son idénticas, los espacios fila de A y B son iguales, de modo que $U = W$.

5.6. DEPENDENCIA E INDEPENDENCIA LINEAL

A continuación se definen las nociones de dependencia e independencia lineal. Estos conceptos juegan un papel esencial dentro de la teoría del álgebra lineal y de las matemáticas en general.

Definición: Sea V un espacio vectorial sobre un cuerpo K. Se dice que los vectores $v_1, \ldots, v_m \in V$ son *linealmente dependientes* sobre K, o simplemente *dependientes*, si existen escalares $a_1, \ldots, a_m \in K$, no todos 0, tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0 \qquad (*)$$

En caso contrario se dice que los vectores son *linealmente independientes* sobre K, o simplemente *independientes*.

Observemos que la relación (*) se verificará siempre si los a_i son todos 0. Si la relación sólo se verifica en este caso, es decir,

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0 \quad \text{implica} \quad a_1 = 0, \ldots, a_m = 0$$

los vectores serán linealmente independientes. Sin embargo, si (*) también es válida cuando uno de los a_i no es 0, los vectores serán linealmente dependientes.

Se dice que un conjunto $\{v_1, v_2, \ldots, v_m\}$ de vectores es linealmente dependiente o independiente según lo sean los vectores $v_1, v_2, \ldots, v_m$. Un conjunto infinito S de vectores es linealmente dependiente si existen vectores $u_1, \ldots, u_k$ en S que lo son; en caso contrario, S es linealmente independiente.

Las siguientes observaciones derivan de las definiciones precedentes.

Nota 1: Si 0 es uno de los vectores $v_1, \ldots, v_m$, digamos $v_1 = 0$, los vectores deben ser linealmente dependientes, ya que

$$1v_1 + 0v_2 + \cdots + 0v_m = 1 \cdot 0 + 0 + \cdots + 0 = 0$$

y el coeficiente de v_1 es distinto de 0.

Nota 2: Cualquier vector no nulo v es por sí solo linealmente independiente, debido a que

$$kv = 0, \quad v \neq 0 \quad \text{implica} \quad k = 0$$

Nota 3: Si dos de los vectores $v_1, v_2, \ldots, v_m$ son iguales, o si uno es un múltiplo escalar de otro, digamos $v_1 = kv_2$, los vectores son linealmente dependientes, puesto que

$$v_1 - kv_2 + 0v_3 + \cdots + 0v_m = 0$$

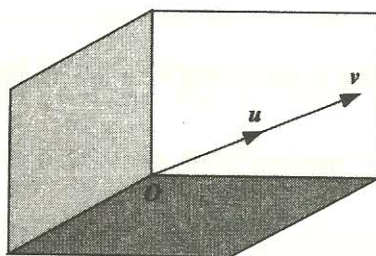
y el coeficiente de v_1 no es 0.

Nota 4: Dos vectores v_1 y v_2 son linealmente dependientes si y sólo si uno de ellos es múltiplo del otro.

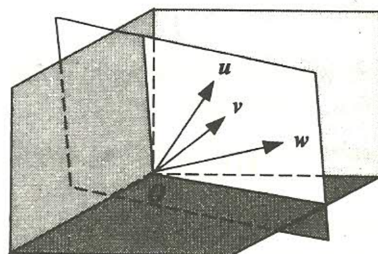
Nota 5: Si el conjunto $\{v_1, \ldots, v_m\}$ es linealmente independiente, cualquier reordenación de los vectores $\{v_{i_1}, v_{i_2}, \ldots, v_{i_m}\}$ también lo es.

Nota 6: Si un conjunto S de vectores es linealmente independiente, necesariamente lo es cualquier subconjunto de S. Alternativamente, si S contiene un subconjunto linealmente dependiente, S es linealmente dependiente.

Nota 7: En el espacio real $\mathbf{R}^3$, la dependencia lineal de vectores puede describirse geométricamente en los siguientes términos: *a*) Dos vectores cualesquiera u y v son linealmente dependientes si y sólo si yacen en la misma recta que pasa por el origen, como se ilustra en la Figura 5-2(*a*). *b*) Tres vectores cualesquiera u, v y w son linealmente dependientes si y sólo si yacen en el mismo plano que pasa por el origen, como se muestra en la Figura 5-2(*b*).



(*a*) u y v son linealmente dependientes



(*b*) u, v y w son linealmente dependientes

Figura 5-2.

A continuación se dan otros ejemplos de vectores linealmente dependientes e independientes.

EJEMPLO 5.5

a) Los vectores $u = (1, -1, 0)$, $v = (1, 3, -1)$ y $w = (5, 3, -2)$ son linealmente dependientes porque

$$3(1, -1, 0) + 2(1, 3, -1) - (5, 3, -2) = (0, 0, 0)$$

Esto es, $3u + 2v - w = 0$.

b) Probamos que los vectores $u = (6, 2, 3, 4)$, $v = (0, 5, -3, 1)$ y $w = (0, 0, 7, -2)$ son linealmente independientes. Supongamos que $xu + yv + zw = 0$, donde x, y y z son escalares desconocidos. En tal caso,

$$\begin{aligned}(0, 0, 0, 0) &= x(6, 2, 3, 4) + y(0, 5, -3, 1) + z(0, 0, 7, -2) = \\ &= (6x, 2x + 5y, 3x - 3y + 7z, 4x + y - 2z)\end{aligned}$$

y así, por la igualdad de las componentes correspondientes,

$$\begin{aligned}6x \qquad\qquad &= 0\\ 2x + 5y \qquad &= 0\\ 3x - 3y + 7z &= 0\\ 4x + y - 2z &= 0\end{aligned}$$

La primera ecuación conduce a $x = 0$; la segunda, con $x = 0$, proporciona $y = 0$, y la tercera, con $x = 0$ e $y = 0$, lleva a $z = 0$. De este modo,

$$xu + yv + zw = 0 \qquad \text{implica} \qquad x = 0,\ y = 0,\ z = 0$$

En consecuencia, u, v y w son linealmente independientes.

COMBINACIONES LINEALES Y DEPENDENCIA LINEAL

Las nociones de combinación lineal y dependencia lineal están estrechamente relacionadas. De forma específica, demostraremos que los vectores $v_1, v_2, \ldots, v_m$, cuando hay más de uno, son linealmente dependientes si y sólo si uno de ellos es una combinación lineal de los otros.

Supongamos, por ejemplo, que v_i es una combinación lineal de los otros:

$$v_i = a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_m v_m$$

Entonces, sumando $-v_i$ a ambos miembros, obtenemos

$$a_1 v_1 + \cdots + a_{i-1} v_{i-1} - v_i + a_{i+1} v_{i+1} + \cdots + a_m v_m = 0$$

donde el coeficiente de v_i no es 0; por consiguiente, los vectores son linealmente dependientes. Recíprocamente, supongamos que los vectores son linealmente dependientes, digamos

$$b_1 v_1 + \cdots + b_j v_j + \cdots + b_m v_m = 0 \qquad \text{donde} \qquad b_j \neq 0$$

En ese caso,

$$v_j = -b_j^{-1} b_1 v_1 - \cdots - b_j^{-1} b_{j-1} v_{j-1} - b_j^{-1} b_{j+1} v_{j+1} - \cdots - b_j^{-1} b_m v_m$$

de modo que v_j es una combinación lineal del resto de los vectores.

Establecemos ahora un resultado algo más potente que el anterior (véase el Problema 5.36 para su demostración), que tiene numerosas consecuencias importantes.

Lema 5.10: Supongamos que dos o más vectores no nulos $v_1, v_2, \ldots, v_m$ son linealmente dependientes. Entonces uno de los vectores es una combinación lineal de los precedentes, esto es, existe un $k > 1$ tal que

$$v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1}$$

EJEMPLO 5.6. Consideremos la siguiente matriz en forma escalonada:

$$A = \begin{pmatrix} 0 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 4 & -4 & 4 & -4 & 4 \\ 0 & 0 & 0 & 0 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & 6 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Observemos que las filas R_2, R_3 y R_4 tienen ceros en la segunda columna (bajo el elemento pivote de R_1) y que, por tanto, cualquier combinación lineal de R_2, R_3 y R_4 debe tener un cero como segunda componente. Siendo así, R_1 no puede ser una combinación lineal de las filas no nulas situadas bajo ella. De forma similar, las filas R_3 y R_4 tienen ceros en la tercera columna, bajo el elemento pivote de R_2; por consiguiente, R_2 no puede ser una combinación lineal de las filas no nulas situadas bajo ella. Finalmente, R_3 no puede ser un múltiplo de R_4, puesto que ésta tiene un cero en la quinta columna, bajo el elemento pivote de aquélla. Mirando las filas no nulas de abajo a arriba, R_4, R_3, R_2, R_1, constatamos que ninguna de ellas es combinación lineal de las precedentes. De este modo, las filas no nulas son linealmente independientes, de acuerdo con el Lema 5.10.

El argumento del ejemplo anterior puede aplicarse a las filas no nulas de cualquier matriz escalonada. Así pues, tenemos el siguiente resultado (demostrado en el Problema 5.37), muy útil.

Teorema 5.11: Las filas no nulas de una matriz en forma escalonada son linealmente independientes.

5.7. BASES Y DIMENSION

Comenzamos estableciendo dos caminos equivalentes (Problema 5.30) para definir una base de un espacio vectorial V.

Definición A: Un conjunto $S = \{u_1, u_2, \ldots, u_n\}$ de vectores es una *base* de V si se verifican las dos condiciones:

1. $u_1, u_2, \ldots, u_n$ son linealmente independientes.
2. $u_1, u_2, \ldots, u_n$ generan V.

Definición B: Un conjunto $S = \{u_1, u_2, \ldots, u_n\}$ de vectores es una *base* de V si todo vector $v \in V$ puede escribirse de forma única como combinación lineal de sus vectores.

Se dice que un espacio vectorial V es de *dimensión finita n* o que es *n-dimensional*, escrito

$$\dim V = n$$

si V tiene una base como la anterior, con n elementos. La dimensión está bien definida, a la vista del siguiente teorema (demostrado en el Problema 5.40).

Teorema 5.12: Sea V un espacio vectorial de dimensión finita. Entonces todas las bases de V tienen el mismo número de elementos.

El espacio vectorial $\{0\}$ tiene dimensión 0, por definición. Cuando un espacio vectorial no es de dimensión finita, se dice que es de *dimensión infinita*.

EJEMPLO 5.7

a) Consideremos el espacio vectorial $\mathbf{M}_{2,3}$ de todas las matrices 2×3 sobre un cuerpo K. Las seis matrices siguientes forman una base de $\mathbf{M}_{2,3}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Con mayor generalidad, en el espacio vectorial $\mathbf{M}_{r,s}$ de las matrices $r \times s$, sea E_{ij} la matriz cuya entrada ij es 1, siendo 0 las restantes. Todas las matrices E_{ij} tales constituyen una base de $\mathbf{M}_{r,s}$, denominada su *base usual*. Consecuentemente, $\dim \mathbf{M}_{r,s} = rs$. En particular, $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, 0, \ldots, 0, 1)$ forman la base usual de K^n.

b) Consideremos el espacio vectorial $P_n(t)$ de los polinomios de grado $\leq n$. Los polinomios $1, t, t^2, \ldots, t^n$ forman una base de $\mathbf{P}_n(t)$ y por tanto $\dim \mathbf{P}_n(t) = n + 1$.

El teorema fundamental anterior sobre dimensión es una consecuencia del importante «lema de sustitución» (demostrado en el Problema 5.39) que sigue:

Lema 5.13: Supongamos que $\{v_1, v_2, \ldots, v_n\}$ genera V y que $\{w_1, w_2, \ldots, w_m\}$ es linealmente independiente. En ese caso, $m \leq n$ y V está generado por un conjunto de la forma

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Así, en particular, $n + 1$ o más vectores en V son linealmente dependientes.

Observemos, en el lema precedente, que hemos sustituido m vectores del conjunto generador por los m vectores independientes y aún conservamos un conjunto generador.

Los teoremas enunciados a continuación (y demostrados en los Problemas 5.41, 5.42 y 5.43, respectivamente) se utilizarán con frecuencia.

Teorema 5.14: Sea V un espacio vectorial de dimensión finita n.

i) $n + 1$ o más vectores en V son linealmente dependientes.

ii) Todo conjunto linealmente independiente $S = \{u_1, u_2, \ldots, u_n\}$, con n elementos, es una base de V.

iii) Todo conjunto generador $T = \{v_1, v_2, \ldots, v_n\}$ de V, con n elementos, es una base de V.

Teorema 5.15: Supongamos que S genera un espacio vectorial V.

i) Cualquier número máximo de vectores linealmente independientes en S es una base de V.

ii) Si se suprime de S todo vector que sea combinación lineal de los precedentes, los vectores que quedan constituyen una base de V.

Teorema 5.16: Sean V un espacio vectorial de dimensión finita y $S = \{u_1, u_2, \ldots, u_r\}$ un conjunto de vectores linealmente independientes en V. En ese caso, S es parte de una base de V, es decir, S puede extenderse a una base de V.

EJEMPLO 5.8

a) Consideremos en $\mathbf{R}^4$ los cuatro vectores:

$$(1, 1, 1, 1) \qquad (0, 1, 1, 1) \qquad (0, 0, 1, 1) \qquad (0, 0, 0, 1)$$

Nótese que los vectores formarán una matriz escalonada, por lo que son linealmente independientes. Más aún, dado que $\dim \mathbf{R}^4 = 4$, los vectores constituyen una base de $\mathbf{R}^4$.

b) Consideremos los $n + 1$ polinomios en $\mathbf{P}_n(t)$:

$$1,\ t - 1,\ (t-1)^2,\ \ldots,\ (t-1)^n$$

El grado de $(t-1)^k$ es k, luego ningún polinomio puede ser combinación lineal de los precedentes. Además, constituyen una base de $\mathbf{P}_n(t)$ porque $\dim \mathbf{P}_n(t) = n + 1$.

DIMENSION Y SUBESPACIOS

El siguiente teorema (demostrado en el Problema 5.44) nos da la relación básica entre la dimensión de un espacio vectorial y la de un subespacio.

Teorema 5.17: Sea W un subespacio de un espacio n-dimensional V. Entonces $\dim W \leq n$. Si, en particular, $\dim W = n$, necesariamente $W = V$.

EJEMPLO 5.9. Sea W un subespacio del espacio real $\mathbf{R}^3$. Tenemos $\dim \mathbf{R}^3 = 3$; por consiguiente, según el Teorema 5.17, la dimensión de W sólo puede ser 0, 1, 2 ó 3. Podemos distinguir los casos:

i) $\dim W = 0$, con lo que $W = \{0\}$, un punto.
ii) $\dim W = 1$, con lo que W es una recta por el origen.
iii) $\dim W = 2$, con lo que W es un plano por el origen.
iv) $\dim W = 3$, con lo que W es el espacio $\mathbf{R}^3$ entero.

RANGO DE UNA MATRIZ

Sea A una matriz $m \times n$ arbitraria sobre un cuerpo K. Recordemos que el espacio fila de A es el subespacio de K^n generado por sus filas y que el espacio columna de A es el subespacio de K^m generado por sus columnas.

El *rango por filas* de una matriz A es igual al número máximo de filas linealmente independientes o, equivalentemente, a la dimensión del espacio fila de A. Análogamente, el *rango por columnas* de A es igual al número máximo de columnas linealmente independientes o, equivalentemente, a la dimensión del espacio columna de A.

Aunque f-lin A es un subespacio de K^n y c-lin A uno de K^m, donde n puede no coincidir con m, tenemos el importante resultado (demostrado en el Problema 5.53) que sigue.

Teorema 5.18: Los rangos por filas y por columnas de cualquier matriz A son iguales.

Definición: El *rango* de la matriz A, escrito rango A, es el valor común de su rango por filas y su rango por columnas.

El rango de una matriz puede calcularse fácilmente usando reducción por filas, como se ilustra en el siguiente ejemplo.

EJEMPLO 5.10. Supongamos que queremos hallar una base y la dimensión del espacio fila de

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{pmatrix}$$

Reducimos A a forma escalonada utilizando las operaciones elementales entre filas:

$$A \sim \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & -6 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Recordemos que las matrices equivalentes por filas tienen el mismo espacio fila. De este modo, las filas no nulas de la matriz escalonada, que son independientes de acuerdo con el Teorema 5.11, forman una base del espacio fila de A. Siendo así, dim f-lin $A = 2$ y por tanto rango $A = 2$.

5.8. ECUACIONES LINEALES Y ESPACIOS VECTORIALES

Consideremos un sistema de m ecuaciones lineales con n incógnitas $x_1, \dots, x_n$ sobre un cuerpo K:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad [5.1]$$

o la ecuación matricial equivalente

$$AX = B$$

donde $A = (a_{ij})$ es la matriz de los coeficientes y $X = (x_i)$ y $B = (b_i)$ los vectores columna constituidos por las incógnitas y las constantes, respectivamente. Recordemos que la *matriz ampliada* del sistema se define como la matriz

$$(A, B) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Nota 1: Se dice que las ecuaciones lineales [5.1] son dependientes o independientes según lo sean los vectores correspondientes, esto es, las filas de la matriz ampliada.

Nota 2: Dos sistemas de ecuaciones lineales son equivalentes si y sólo si las matrices ampliadas correspondientes son equivalentes por filas, es decir, tienen el mismo espacio fila.

Nota 3: Siempre podemos reemplazar un sistema de ecuaciones por uno de ecuaciones independientes, como por ejemplo un sistema en forma escalonada. El número de ecuaciones independientes será siempre igual al rango de la matriz ampliada.

Observemos que el sistema [5.1] es también equivalente a la ecuación vectorial

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}$$

El comentario anterior nos conduce al teorema de existencia básico, enunciado a continuación.

Teorema 5.19: Las tres afirmaciones siguientes son equivalentes.

a) El sistema de ecuaciones lineales $AX = B$ tiene alguna solución.

b) B es combinación lineal de las columnas de A.

c) La matriz de coeficientes A y la ampliada (A, B) tienen el mismo rango.

Recuérdese que una matriz $m \times n$ A puede verse como una función $A: K^n \to K^m$. Así un vector B pertenecerá a la imagen de A si y sólo si la ecuación $AX = B$ tiene una solución. Esto significa que la imagen (recorrido) de la función A, escrito Im A, es precisamente el espacio columna de A. En consecuencia,

$$\dim (\text{Im } A) = \dim (\text{c-lin } A) = \text{rango } A$$

Utilizaremos este hecho para probar (Problema 5.59) el siguiente resultado básico relativo a sistemas de ecuaciones lineales homogéneos.

Teorema 5.20: La dimensión del espacio solución W del sistema de ecuaciones lineales homogéneo $AX = 0$ es $n - r$, donde n es el número de incógnitas y r el rango de la matriz de los coeficientes A.

En caso de que el sistema $AX = 0$ esté en forma escalonada, éste tendrá precisamente $n - r$ variables libres, digamos $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}}$. Sea v_j la solución obtenida haciendo $x_{i_j} = 1$ (o cualquier constante no nula) y el resto de las variables libres iguales a 0. Entonces las soluciones $v_1, \ldots, v_{n-r}$ serán linealmente independientes (Problema 5.58) y formarán una base del espacio solución.

EJEMPLO 5.11. Supongamos que queremos hallar la dimensión y una base del espacio solución, W, del sistema:

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ x + 2y + 3z + s + t &= 0 \\ 3x + 6y + 8z + s + 5t &= 0 \end{aligned}$$

Primero reducimos el sistema a forma escalonada:

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \\ 2z + 4s - 4t &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \end{aligned}$$

El sistema en forma escalonada tiene dos ecuaciones (no nulas) y cinco incógnitas y por tanto tiene $5 - 2 = 3$ variables libres, que son y, s y t. De este modo, $\dim W = 3$. Para obtener una base de W tomemos:

i) $y = 1, s = 0, t = 0$ para llegar a la solución $v_1 = (-2, 1, 0, 0, 0)$.

ii) $y = 0, s = 1, t = 0$ para llegar a la solución $v_2 = (5, 0, -2, 1, 0)$.

iii) $y = 0, s = 0, t = 1$ para llegar a la solución $v_3 = (-7, 0, 2, 0, 1)$.

El conjunto $\{v_1, v_2, v_3\}$ es una base del espacio solución W.

DOS ALGORITMOS PARA HALLAR BASES

Supongamos que se nos dan los vectores $u_1, u_2, \ldots, u_r$ en K^n. Sea

$$W = \text{lin}\,(u_1, u_2, \ldots, u_r)$$

el subespacio de K^n generado por dichos vectores. Cada uno de los algoritmos expuestos abajo determinan una base (y por ende la dimensión) de W.

Algoritmo 5.8A (algoritmo del espacio fila)

Paso 1. Construir una matriz A cuyas *filas* sean los vectores dados.

Paso 2. Reducir por filas A a forma escalonada.

Paso 3. Extraer las filas no nulas de la matriz escalonada.

El algoritmo precedente ya apareció en esencia en el Ejemplo 5.10. El próximo se ilustrará en el Ejemplo 5.12 y utiliza el resultado referente a sistemas de ecuaciones lineales inhomogéneos que se dio con anterioridad.

Algoritmo 5.8B (algoritmo de expulsión)

Paso 1. Construir una matriz M cuyas *columnas* sean los vectores dados.

Paso 2. Reducir por filas M a forma escalonada.

Paso 3. Para cada columna C_k sin pivote en la matriz escalonada, suprimir (expulsar) el vector v_k del conjunto.

Paso 4. Extraer los restantes vectores (correspondientes a columnas con pivote).

EJEMPLO 5.12. Sea W el subespacio de $\mathbf{R}^5$ generado por los vectores:

$$v_1 = (1, 2, 1, -2, 3) \qquad v_2 = (2, 5, -1, 3, -2) \qquad v_3 = (1, 3, -2, 5, -5)$$
$$v_4 = (3, 1, 2, -4, 1) \qquad v_5 = (5, 6, 1, -1, -1)$$

Usamos el Algoritmo *5.8B* para hallar la dimensión y una base de W.

Comenzamos por construir la matriz M, cuyas columnas son los vectores dados, y reducirla a forma escalonada:

$$M = \begin{pmatrix} 1 & 2 & 1 & 3 & 5 \\ 2 & 5 & 3 & 1 & 6 \\ 1 & -1 & -2 & 2 & 1 \\ -2 & 3 & 5 & -4 & -1 \\ 3 & -2 & -5 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 5 \\ 0 & 1 & 1 & -5 & -4 \\ 0 & -3 & -3 & -1 & -4 \\ 0 & 7 & 7 & 2 & 9 \\ 0 & -8 & -8 & -8 & -16 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 3 & 5 \\ 0 & 1 & 1 & 7 & -4 \\ 0 & 0 & 0 & -16 & -16 \\ 0 & 0 & 0 & 37 & 37 \\ 0 & 0 & 0 & -48 & -48 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 & 5 \\ 0 & 1 & 1 & 7 & -4 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Observemos que los pivotes de la matriz escalonada aparecen en las columnas C_1, C_2 y C_4. El hecho de que la columna C_3 no tenga pivote significa que el sistema $xv_1 + yv_2 = v_3$ tiene una solución y por consiguiente v_3 es combinación lineal de v_1 y v_2. De forma similar, el hecho de que C_5 no tenga pivote significa que v_5 es combinación lineal de los vectores precedentes. De acuerdo con esto, los vectores v_1, v_2 y v_4, correspondientes a las columnas con pivote en la matriz, constituyen una base de W y $\dim W = 3$.

5.9. SUMAS Y SUMAS DIRECTAS

Sean U y W subconjuntos de un espacio vectorial V. La suma de U y W, escrito $U + W$, consiste en todas las sumas $u + w$ en las que $u \in U$ y $w \in W$. Esto es,

$$U + W = \{u + w\colon\ u \in U,\ w \in W\}$$

Supongamos ahora que U y W son subespacios de V. Nótese que $0 = 0 + 0 \in U + W$, ya que $0 \in U$ y $0 \in W$. Supongamos, además, que $u + w$ y $u' + w'$ pertenecen a $U + W$, con $u, u' \in U$ y $w, w' \in W$. Entonces

$$(u + w) + (u' + w') = (u + u') + (w + w') \in U + W$$

y, para todo escalar k,

$$k(u + w) = ku + kw \in U + W$$

Así hemos demostrado el siguiente teorema.

Teorema 5.21: La suma $U + W$ de los subespacios U y W de V es también un subespacio de V.

Recordemos que la intersección $U \cap W$ es subespacio de V. El teorema enunciado a continuación, y demostrado en el Problema 5.69, relaciona las dimensiones de estos subespacios.

Teorema 5.22: Sean U y W subespacios de dimensión finita de un espacio vectorial V. En ese caso, $U + W$ tiene dimensión finita y

$$\dim (U + W) = \dim U + \dim W - \dim (U \cap W)$$

EJEMPLO 5.13. Supongamos que U y W son los planos xy e yz en $\mathbf{R}^3$, respectivamente. Es decir,

$$U = \{(a, b, 0)\} \qquad \text{y} \qquad W = \{(0, b, c)\}$$

Nótese que $\mathbf{R}^3 = U + W$; por consiguiente, $\dim (U + W) = 3$. Asimismo, $\dim U = 2$ y $\dim W = 2$. Según el Teorema 5.22,

$$3 = 2 + 2 - \dim (U \cap W) \qquad \text{o} \qquad \dim (U \cap W) = 1$$

Esto concuerda con los hechos de que $U \cap W$ es el eje y (Fig. 5-3) y que dicho eje tiene dimensión 1.

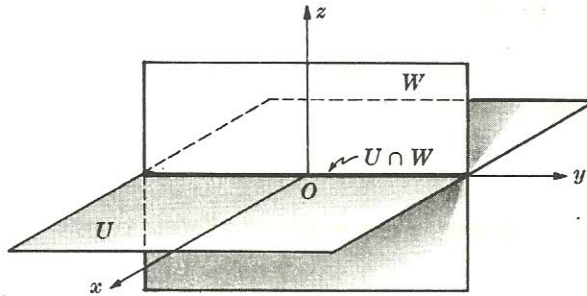


Figura 5-3.

SUMAS DIRECTAS

Se dice que el espacio vectorial V es la *suma directa* de sus subespacios U y W, denotado por

$$V = U \oplus W$$

si todo vector $v \in V$ puede escribirse de una y sólo una forma como $v = u + w$, con $u \in U$ y $w \in W$.

El teorema que sigue, demostrado en el Problema 5.70, caracteriza tal descomposición.

Teorema 5.23: El espacio vectorial V es la suma directa de sus subespacios U y W si y sólo si i) $V = U + W$ y ii) $U \cap W = \{0\}$.

EJEMPLO 5.14

a) En el espacio vectorial $\mathbf{R}^3$, sean U el plano xy y W el yz:

$$U = \{(a, b, 0)\colon a, b \in \mathbf{R}\} \quad \text{y} \quad W = \{(0, b, c)\colon b, c \in \mathbf{R}\}$$

Entonces $\mathbf{R}^3 = U + W$, puesto que todo vector en $\mathbf{R}^3$ es la suma de un vector de U y uno de W. Sin embargo, $\mathbf{R}^3$ no es la suma directa de U y W, ya que dichas sumas no son únicas; por ejemplo,

$$(3, 5, 7) = (3, 1, 0) + (0, 4, 7) \quad \text{y también} \quad (3, 5, 7) = (3, -4, 0) + (0, 9, 7)$$

b) En $\mathbf{R}^3$, sean U el plano xy y W el eje z:

$$U = \{(a, b, 0)\colon a, b \in \mathbf{R}\} \quad \text{y} \quad W = \{(0, 0, c)\colon c \in \mathbf{R}\}$$

Ahora cualquier vector $(a, b, c) \in \mathbf{R}^3$ puede escribirse como la suma de un vector de U y uno de W de una y sólo una forma:

$$(a, b, c) = (a, b, 0) + (0, 0, c)$$

En consecuencia, $\mathbf{R}^3$ es la suma directa de U y W, o sea, $\mathbf{R}^3 = U \oplus W$. Alternativamente, $\mathbf{R}^3 = U \oplus W$ puesto que $\mathbf{R}^3 = U + W$ y $U \cap W = \{0\}$.

SUMAS DIRECTAS GENERALES

La noción de suma directa se extiende a más de un sumando de la manera obvia. Esto es, V es la *suma directa* de los subespacios $W_1, W_2, \ldots, W_r$, escrito

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

si todo vector $v \in V$ puede escribirse de una y sólo una forma como

$$v = w_1 + w_2 + \cdots + w_r$$

donde $w_1 \in W_1$, $w_2 \in W_2$, ..., $w_r \in W_r$.

Son aplicables los siguientes teoremas.

Teorema 5.24: Sea $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$. Además, para cada i, supongamos que S_i es un subconjunto linealmente independiente de W_i. En tal caso,

a) La unión $S = \bigcup_i S_i$ es linealmente independiente en V.

b) Si S_i es una base de W_i, entonces $S = \bigcup_i S_i$ es una base de V.

c) $\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_r$.

Teorema 5.25: Supongamos que $V = W_1 + W_2 + \cdots + W_r$ (donde V tiene dimensión finita) y que

$$\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_r$$

Entonces $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

5.10. COORDENADAS

Sean V un espacio vectorial n-dimensional sobre un cuerpo K y

$$S = \{u_1, u_2, \ldots, u_n\}$$

una base de V. Cualquier vector $v \in V$ puede expresarse de forma única como combinación lineal de los vectores de la base en S, digamos

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

Estos n escalares $a_1, a_2, \ldots, a_n$ se denominan las *coordenadas* de v relativas a la base S, y forman la n-pla $[a_1, a_2, \ldots, a_n]$ en K^n, llamada el *vector coordenado* de v relativo a S. Denotamos este vector por $[v]_S$, o simplemente $[v]$, cuando S viene dada implícitamente. Así

$$[v]_S = [a_1, a_2, \ldots, a_n]$$

Obsérvese que se utilizan corchetes [...], y no paréntesis (...), para designar el vector coordenado.

EJEMPLO 5.15

a) Consideremos el espacio vectorial $\mathbf{P}_2(t)$ de los polinomios de grado ≤ 2. Los polinomios

$$p_1 = 1 \qquad p_2 = t - 1 \qquad p_3 = (t-1)^2 = t^2 - 2t + 1$$

forman una base S de $\mathbf{P}_2(t)$. Sea $v = 2t^2 - 5t + 6$. El vector coordenado de v relativo a la base S se obtiene como sigue.

Tomemos $v = xp_1 + yp_2 + zp_3$ usando los escalares desconocidos x, y, z y simplifiquemos:

$$\begin{aligned} 2t^2 - 5t + 6 &= x(1) + y(t-1) + z(t^2 - 2t + 1) = \\ &= x + yt - y + zt^2 - 2zt + z = \\ &= zt^2 + (y - 2z)t + (x - y + z) \end{aligned}$$

A continuación igualemos entre sí los coeficientes de las mismas potencias de t:

$$\begin{aligned} x - y + z &= 6 \\ y - 2z &= -5 \\ z &= 2 \end{aligned}$$

La solución del sistema anterior es: $x = 3$, $y = -1$, $z = 2$. De este modo,

$$v = 3p_1 - p_2 + 2p_3 \qquad \text{y por tanto} \qquad [v] = [3, -1, 2]$$

b) Consideremos el espacio real $\mathbf{R}^3$. Los vectores

$$u_1 = (1, -1, 0) \qquad u_2 = (1, 1, 0) \qquad u_3 = (0, 1, 1)$$

forman una base S de $\mathbf{R}^3$. Sea $v = (5, 3, 4)$. Las coordenadas de v relativas a la base S se obtienen como sigue.

Tomemos $v = xu_1 + yu_2 + zu_3$, es decir, expresemos v como una combinación lineal de los vectores de la base usando escalares desconocidos x, y, z:

$$\begin{aligned} (5, 3, 4) &= x(1, -1, 0) + y(1, 1, 0) + z(0, 1, 1) = \\ &= (x, -x, 0) + (y, y, 0) + (0, z, z) = \\ &= (x + y, -x + y + z, z) \end{aligned}$$

Después igualemos entre sí las componentes correspondientes para llegar al sistema de ecuaciones lineales equivalentes:

$$x + y = 5 \qquad -x + y + z = 3 \qquad z = 4$$

La solución del sistema es $x = 3$, $y = 2$, $z = 4$. Así

$$v = 3u_1 + 2u_2 + 4u_3 \qquad \text{y por tanto} \qquad [v]_S = [3, 2, 4]$$

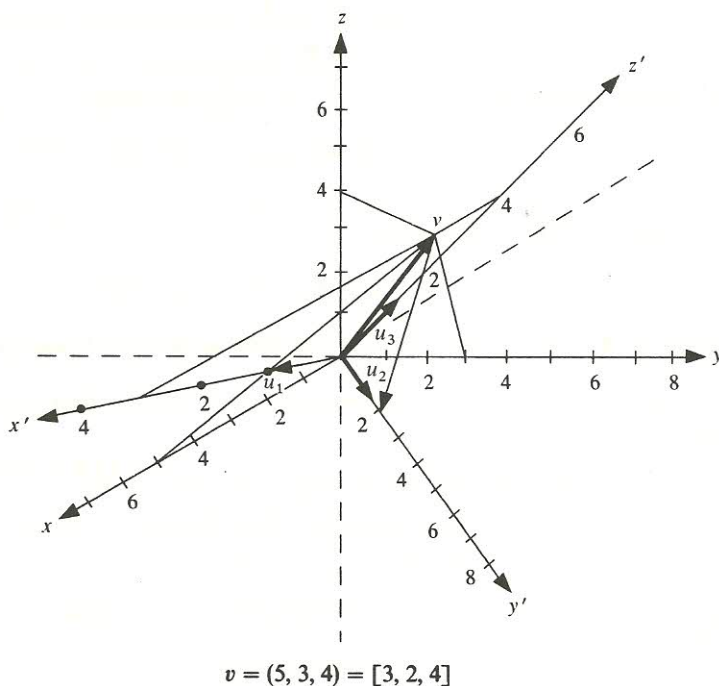
Nota: Existe una interpretación geométrica de las coordenadas de un vector v relativas a una base S del espacio real $\mathbf{R}^n$. Ilustramos esto utilizando la base de $\mathbf{R}^3$, que aparece en el Ejemplo 5.15:

$$S = \{u_1 = (1, -1, 0),\ u_2 = (1, 1, 0),\ u_3 = (0, 1, 1)\}$$

Consideremos primero el espacio $\mathbf{R}^3$ con los ejes x, y, z usuales. En ese caso, los vectores de la base determinan un nuevo sistema de coordenadas de $\mathbf{R}^3$, digamos con ejes x', y', z', tal y como se muestra en la Figura 5-4. O sea:

1. El eje x' está en la dirección de u_1.
2. El eje y' está en la dirección de u_2.
3. El eje z' está en la dirección de u_3.

Además, la unidad de longitud de cada uno de los ejes será igual, respectivamente, a la longitud del vector de la base correspondiente. Así pues, cada vector $v = (a, b, c)$ o, equivalentemente, cada punto $P(a, b, c)$ en $\mathbf{R}^3$ tendrá unas nuevas coordenadas con respecto a los nuevos ejes x', y', z'. Dichas coordenadas son precisamente las de v con respecto a la base S.



$v = (5, 3, 4) = [3, 2, 4]$

Figura 5-4.

ISOMORFISMO ENTRE V Y K^n

Consideremos una base $S = \{u_1, u_2, \ldots, u_n\}$ de un espacio vectorial V sobre un cuerpo K. Hemos probado con anterioridad que a cada vector $v \in V$ le corresponde una única n-pla $[v]_S$ en K^n. Por otra parte, a cada n-pla $[c_1, c_2, \ldots, c_n] \in K^n$ le corresponderá el vector $c_1u_1 + c_2u_2 + \cdots + c_nu_n$

en V. Siendo así, la base S induce una correspondencia uno-a-uno entre los vectores de V y las n-plas de K^n. Más aún, supongamos

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n \qquad \text{y} \qquad w = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

Entonces

$$\begin{aligned} v + w &= (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_n + b_n)u_n \\ kv &= (ka_1)u_1 + (ka_2)u_2 + \cdots + (ka_n)u_n \end{aligned}$$

siendo k un escalar. De acuerdo con ello,

$$[v + w]_S = [a_1 + b_1, \ldots, a_n + b_n] = [a_1, \ldots, a_n] + [b_1, \ldots, b_n] = [v]_S + [w]_S$$

y

$$[kv]_S = [ka_1, ka_2, \ldots, ka_n] = k[a_1, a_2, \ldots, ka_n] = k[v]_S$$

De este modo, la correspondencia uno-a-uno anterior, entre V y K^n, conserva las operaciones de suma y producto por un escalar del espacio vectorial. Decimos, en tal caso, que V y K^n son *isomorfos*, escrito $V \cong K^n$. Establezcamos formalmente este resultado.

Teorema 5.26. Sea V un espacio vectorial n-dimensional sobre un cuerpo K. Entonces V y K^n son isomorfos.

El próximo ejemplo muestra una aplicación práctica de dicho resultado.

EJEMPLO 5.16. Supongamos que queremos determinar si las siguientes matrices son o no linealmente independientes:

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 & -4 \\ 6 & 5 & 4 \end{pmatrix} \qquad C = \begin{pmatrix} 3 & 8 & -11 \\ 16 & 10 & 9 \end{pmatrix}$$

Los vectores coordenados de las matrices anteriores respecto a la base usual [Ejemplo 5.7 *a*)] de $\mathbf{M}_{2,3}$ son:

$$[A] = (1, 2, -3, 4, 0, 1) \qquad [B] = (1, 3, -4, 6, 5, 4) \qquad [C] = (3, 8, -11, 16, 10, 9)$$

Construimos la matriz M cuyas filas son los vectores coordenados precedentes:

$$M = \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 1 & 3 & -4 & 6 & 5 & 4 \\ 3 & 8 & -11 & 16 & 10 & 9 \end{pmatrix}$$

Reducimos por filas M a forma escalonada:

$$M \sim \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 2 & -2 & 4 & 10 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Como la forma escalonada sólo tiene dos filas no nulas, los vectores coordenados $[A]$, $[B]$ y $[C]$ generan un subespacio de dimensión 2, de modo que son linealmente dependientes. En consecuencia, las matrices iniciales A, B y C son linealmente dependientes.

5.11. CAMBIO DE BASE

En la Sección 5.10 se demostró que es posible representar cada vector de un espacio vectorial V mediante una n-pla, una vez que se ha seleccionado una base S de V. Surge de forma natural la siguiente pregunta: ¿Cómo variará nuestra representación si elegimos otra base? Para responder debemos redefinir algunos términos. Concretamente, supongamos que $a_1, a_2, \ldots, a_n$ son las coordenadas de un vector v, relativas a una base S de V. En ese caso representaremos v por su *vector columna coordenado*, denotado y definido por

$$[v]_S = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix} = (a_1, a_2, \ldots, a_n)^T$$

Subrayamos el hecho de que en esta sección $[v]_S$ es una matriz $n \times 1$ y no simplemente un elemento de K^n. (El significado de $[v]_S$ estará siempre claro en este contexto.)

Supongamos que $S = \{u_1, u_2, \ldots, u_n\}$ es una base de un espacio vectorial V y que $S' = \{v_1, v_2, \ldots, v_n\}$ es otra. Por ser S base, cada vector de S' puede escribirse de forma única como combinación lineal de los elementos de S. Por ejemplo,

$$\begin{aligned} v_1 &= c_{11}u_1 + c_{12}u_2 + \cdots + c_{1n}u_n \\ v_2 &= c_{21}u_1 + c_{22}u_2 + \cdots + c_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_n &= c_{n1}u_1 + c_{n2}u_2 + \cdots + c_{nn}u_n \end{aligned}$$

Denotemos por P la traspuesta de la matriz de coeficientes anterior:

$$P = \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{pmatrix}$$

O sea, $P = (p_{ij})$, donde $p_{ij} = c_{ji}$. Esta P recibe el nombre de *matriz de cambio de base* (o el de *matriz de transición*) desde la «antigua base» S hasta la «nueva base» S'.

Nota: Como los vectores $v_1, v_2, \ldots, v_n$ de S' son linealmente independientes, la matriz P es invertible (Problema 5.84). De hecho (Problema 5.80), su inversa P^{-1} es la matriz de cambio de base desde la S', volviendo a la S.

EJEMPLO 5.17. Consideremos las dos bases de $\mathbf{R}^2$:

$$S = \{u_1 = (1, 2), u_2 = (3, 5)\} \qquad E = \{e_1 = (1, 0), e_2 = (0, 1)\}$$

De $u_1 = e_1 + 2e_2$, $u_2 = 3e_1 + 5e_2$ se sigue que

$$\begin{aligned} e_1 &= -5u_1 + 2u_2 \\ e_2 &= 3u_1 - u_2 \end{aligned}$$

Escribir los coeficientes de u_1 y u_2 como columnas proporciona la matriz de cambio de base P desde la base S hasta la base usual E:

$$P = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$$

Además, como E es la base usual,

$$u_1 = (1, 2) = e_1 + 2e_2$$
$$u_2 = (3, 5) = 3e_1 + 5e_2$$

Escribir los coeficientes de e_1 y e_2 como columnas proporciona la matriz de cambio de base Q desde la base E, regresando a la base S:

$$Q = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$$

Observemos que P y Q son inversas:

$$PQ = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

El siguiente teorema (demostrado en el Problema 5.19) nos describe cómo son afectados los vectores (columna) coordenados por un cambio de base.

Teorema 5.27: Sea P la matriz de cambio de base desde una base S hasta otra S' en un espacio vectorial V. En ese caso, para todo vector $v \in V$, tenemos

$$P[v]_{S'} = [v]_S \qquad \text{y, por consiguiente,} \qquad P^{-1}[v]_S = [v]_{S'}$$

Nota: A pesar de llamarse P la matriz de cambio de base desde la antigua base S hasta la nueva S', es P^{-1} la que transforma las coordenadas de v relativas a S en las coordenadas de v relativas a S'.

Ilustramos el teorema precedente en el caso dim $V = 3$. Supongamos que P es la matriz de cambio de base desde la base $S = \{u_1, u_2, u_3\}$ hasta la $S' = \{v_1, v_2, v_3\}$; es decir,

$$v_1 = a_1u_1 + a_2u_2 + a_3u_3$$
$$v_2 = b_1u_1 + b_2u_2 + b_3u_3$$
$$v_3 = c_1u_1 + c_2u_2 + c_3u_3$$

Por tanto,

$$P = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Supongamos ahora que $v \in V$ y digamos $v = k_1v_1 + k_2v_2 + k_3v_3$. Expresando lo anterior en términos de v_1, v_2 y v_2 obtenemos

$$v = k_1(a_1u_1 + a_2u_2 + a_3u_3) + k_2(b_1u_1 + b_2u_2 + b_3u_3) + k_3(c_1u_1 + c_2u_2 + c_3u_3) =$$
$$= (a_1k_1 + b_1k_2 + c_1k_3)u_1 + (a_2k_1 + b_2k_2 + c_2k_3)u_2 + (a_3k_1 + b_3k_2 + c_3k_3)u_3$$

Así

$$[v]_{S'} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} \qquad \text{y} \qquad [v]_S = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix}$$

De acuerdo con esto,

$$P[v]_{S'} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix} = [v]_S$$

Asimismo, multiplicando la ecuación precedente por P^{-1},

$$P^{-1}[v]_S = P^{-1}P[v]_{S'} = I[v]_{S'} = [v]_{S'}$$

Nota: Supongamos que $S = \{u_1, u_2, \ldots, u_n\}$ es una base de un espacio vectorial V sobre un cuerpo K y que $P = (p_{ij})$ es una matriz no singular arbitraria sobre K. Entonces los n vectores

$$v_i = p_{1i}u_1 + p_{2i}u_2 + \cdots + p_{ni}u_n \qquad i = 1, 2, \ldots, n$$

son linealmente independientes (Problema 5.84) y por ende forman otra base S' de V. Más aún, P será la matriz de cambio de base desde S hasta la nueva base S'.

PROBLEMAS RESUELTOS

ESPACIOS VECTORIALES

5.1. Demostrar el Teorema 5.1.

i) Por el axioma $[A_2]$, con $u = 0$, sabemos que $0 + 0 = 0$. Por tanto, según el axioma $[M_1]$,

$$k0 = k(0 + 0) = k0 + k0$$

Sumando $-k0$ a ambos miembros conseguimos el resultado deseado.

ii) Una propiedad de K es que $0 + 0 = 0$. Por consiguiente, de acuerdo con el axioma $[M_2]$, $0u = (0 + 0)u = 0u + 0u$. Sumando $-0u$ a ambos miembros llegamos al resultado requerido.

iii) Supongamos que $ku = 0$ y $k \neq 0$. Entonces existe un escalar k^{-1} tal que $k^{-1}k = 1$, luego

$$u = 1u = (k^{-1}k)u = k^{-1}(ku) = k^{-1}0 = 0$$

iv) Utilizando $u + (-u) = 0$ obtenemos $0 = k0 = k(u + (-u)) = ku + k(-u)$. Sumando $-ku$ a ambos miembros, $-ku = k(-u)$.

Usando $k + (-k) = 0$ obtenemos $0 = 0u = (k + (-k))u = ku + (-k)u$. Sumando $-ku$ a ambos miembros, $-ku = (-k)u$. De este modo, $(-k)u = k(-u) = -ku$.

5.2. Probar que para todo escalar k y todo par de vectores u y v, $k(u - v) = ku - kv$.

Utilizamos la definición de resta $u - v \equiv u + (-v)$ y el resultado $k(-v) = -kv$ para obtener:

$$k(u - v) = k(u + (-v)) = ku + k(-v) = ku + (-kv) = ku - kv$$

5.3. Sea V el conjunto de todas las funciones de un conjunto no vacío X en un cuerpo K. Para todo par de funciones f, $g \in V$ y todo escalar $k \in K$, sean $f + g$ y kf las funciones en V definidas como sigue:

$$(f + g)(x) = f(x) + g(x) \qquad \text{y} \qquad (kf)(x) = kf(x) \qquad \forall x \in X$$

Demostrar que V es un espacio vectorial sobre K.

Como X es no vacío, también lo será V. Necesitamos ahora probar que se verifican todos los axiomas de un espacio vectorial.

[A_1] Sean f, g, $h \in V$. Para demostrar que $(f + g) + h = f + (g + h)$ es necesario comprobar que las dos funciones asignan el mismo valor a cada $x \in X$. Ahora bien,

$$((f + g) + h)(x) = (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) \qquad \forall x \in X$$
$$(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) \qquad \forall x \in X$$

Pero $f(x)$, $g(x)$ y $h(x)$ son escalares en el cuerpo K, donde la suma de escalares es asociativa; de aquí

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$$

De acuerdo con esto, $(f + g) + h = f + (g + h)$.

[A_2] Denotemos por $\mathbf{0}$ la función cero: $\mathbf{0}(x) = 0$, $\forall x \in X$. Para cada función $f \in V$,

$$(f + \mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 = f(x) \qquad \forall x \in X$$

Así pues, $f + \mathbf{0} = f$ y $\mathbf{0}$ es el vector cero en V.

[A_3] Para cada función $f \in V$, sea $-f$ la función definida por $(-f)(x) = -f(x)$. Entonces

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0 = \mathbf{0}(x) \qquad \forall x \in X$$

De aquí $f + (-f) = \mathbf{0}$.

[A_4] Sean f, $g \in V$. En tal caso,

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x) \qquad \forall x \in X$$

Por tanto, $f + g = g + f$. [Nótese que $f(x) + g(x) = g(x) + f(x)$ deriva del hecho de que $f(x)$ y $g(x)$ son escalares en el cuerpo K, en el que la suma es conmutativa.]

[M_1] Sean f, $g \in K$. Tendremos

$$(k(f + g))(x) = k((f + g)(x)) = k(f(x) + g(x)) = kf(x) + kg(x) =$$
$$= (kf)(x) + kg)(x) = (kf + kg)(x) \qquad \forall x \in X$$

En consecuencia, $k(f + g) = kf + kg$. [Nótese que $k(f(x) + g(x)) = kf(x) + kg(x)$ se sigue del hecho de que k, $f(x)$ y $g(x)$ son escalares en el cuerpo K, donde el producto es distributivo respecto a la suma.]

[M_2] Sean $f \in V$ y a, $b \in K$. Entonces

$$((a + b)f)(x) = (a + b)f(x) = af(x) + bf(x) = (af)(x) + bf(x) =$$
$$= (af + bf)(x), \qquad \forall x \in X$$

Por consiguiente, $(a + b)f = af + bf$.

$[M_3]$ Sean $f \in V$ y $a, b \in K$. Entonces

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a(bf)(x) = (a(bf))(x) \qquad \forall x \in X$$

Por tanto, $(ab)f = a(bf)$.

$[M_4]$ Sea $f \in V$. Para la unidad $1 \in K$, $(1f)(x) = 1f(x)) = f(x)$, $\forall x \in X$. Así $1f = f$.

Como se satisfacen todos los axiomas, V es un espacio vectorial sobre K.

SUBESPACIOS

5.4. Demostrar el Teorema 5.2.

Supongamos que W satisface i), ii) y iii). Por i), W es no vacío, mientras que por ii) y iii), las operaciones de suma vectorial y producto por un escalar están bien definidas sobre W. Además, los axiomas $[A_1]$, $[A_4]$, $[M_1]$, $[M_2]$, $[M_3]$ y $[M_4]$ se verifican en W, puesto que los vectores de W pertenecen a V. Por tanto, sólo necesitamos probar que $[A_2]$ y $[A_3]$ también se verifican en W. Por i), W es no vacío, digamos $u \in W$. Entonces, por iii), $0u = 0 \in W$ y $v + 0 = v$ para todo $v \in W$, luego W satisface $[A_2]$. Finalmente, si $v \in W$, necesariamente $(-1)v = -v \in W$ y $v + (-v) = 0$; por consiguiente, W es un subespacio de V. Recíprocamente, si W es un subespacio de V, es claro que se verifican i), ii) y iii).

5.5. Demostrar el Corolario 5.3.

Supongamos que W satisface i) y ii). De acuerdo con i), W es no vacío. Además, si $v, w \in W$, entonces, por ii), $v + w = 1v + 1w \in W$; y si $v \in W$ y $k \in K$, por ii), $kv = kv + 0v \in W$. Así, según el Teorema 5.2, W es un subespacio de V.

Recíprocamente, si W es un subespacio de V, claramente se verifican i) y ii) en W.

5.6. Probar que W es un subespacio de $\mathbf{R}^3$, donde $W = \{(a, b, c)\colon a + b + c = 0\}$, esto es, W consiste en aquellos vectores con la propiedad de que la suma de sus componentes es cero.

$0 = (0, 0, 0) \in W$ porque $0 + 0 + 0 = 0$. Supongamos que $v = (a, b, c)$, $w = (a', b', c')$ pertenecen a W, es decir, $a + b + c = 0$ y $a' + b' + c' = 0$. Entonces, para dos escalares cualesquiera k y k',

$$kv + k'w = k(a, b, c) + k'(a', b', c') = (ka, kb, kc) + (k'a', k'b', k'c') = (ka + k'a', kb + k'b', kc + k'c')$$

y, además,

$$(ka + k'a') + (kb + k'b') + (kc + k'c') = k(a + b + c) + k'(a' + b' + c') = k0 + k'0 = 0$$

De este modo, $kv + k'w \in W$ y por tanto W es un subespacio de $\mathbf{R}^3$.

5.7. Sea V el espacio vectorial de todas las matrices cuadradas $n \times n$ sobre un cuerpo K. Mostrar que W es un subespacio de V, donde:

a) W consiste en las matrices simétricas, es decir, todas las matrices $A = (a_{ij})$ para las que $a_{ji} = a_{ij}$.

b) W consiste en todas las matrices que conmutan con una matriz dada T; esto es,

$$W = \{A \in V\colon AT = TA\}$$

a) $0 \in W$, ya que todas las entradas de 0 son 0 y por tanto iguales. Supongamos ahora que $A = (a_{ij})$ y $B = (b_{ij})$ pertenecen a W, o sea, $a_{ji} = a_{ij}$ y $b_{ji} = b_{ij}$. Para todo par de escalares a,

$b \in K$, $aA + bB$ es la matriz cuya entrada ij es $aa_{ij} + bb_{ij}$. Pero $aa_{ji} + bb_{ji} = aa_{ij} + bb_{ij}$. Siendo así, $aA + bB$ es también simétrica, de modo que W es un subespacio de V.

b) $0 \in W$, puesto que $0T = 0 = T0$. Supongamos ahora que A, $B \in W$, es decir, $AT = TA$ y $BT = TB$. Para todo par de escalares a, $b \in K$,

$$(aA + bB)T = (aA)T + (bB)T = a(AT) + b(BT) = a(TA) + b(TB) =$$
$$= T(aA) + T(bB) = T(aA + bB)$$

De este modo, $aA + bB$ conmuta con T, o sea, pertenece a W; por consiguiente, W es un subespacio de V.

5.8. Sea V el espacio vectorial de todas las matrices 2×2 sobre el cuerpo real $\mathbf{R}$. Probar que W no es un subespacio de V si:

a) W consiste en todas las matrices con determinante nulo.

b) W consiste en todas las matrices A para las que $A^2 = A$.

a) $\left[\text{Recuérdese que } \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.\right]$ Las matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ pertenecen a W, pues $\det(A) = 0$ y $\det(B) = 0$. Sin embargo, $A + B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ no pertenece a W debido a que $\det(A + B) = 1$. Por tanto, W no es un subespacio de V.

b) La matriz unidad $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ pertenece a W, ya que

$$I^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Pero $2I = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ no pertenece a W porque

$$(2I)^2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \neq 2I$$

En consecuencia, W no es un subespacio de V.

5.9. Sea V el espacio vectorial de todas las funciones del cuerpo real $\mathbf{R}$ en $\mathbf{R}$. Probar que W es un subespacio de V, donde W consiste en todas las funciones impares, esto es, aquellas funciones f para las que $f(-x) = -f(x)$.

Denotemos por $\mathbf{0}$ la función cero: $\mathbf{0}(x) = 0$ para todo $x \in \mathbf{R}$. $\mathbf{0} \in W$ puesto que $\mathbf{0}(-x) = 0 = -0 = -\mathbf{0}(x)$. Supongamos f, $g \in W$, o sea, $f(-x) = -f(x)$ y $g(-x) = -g(x)$. Para dos números reales cualesquiera a y b,

$$(af + bg)(-x) = af(-x) + bg(-x) = -af(x) - bg(x) = -(af(x) + bg(x)) = -(af + bg)(x)$$

Así pues, $af + bg \in W$ y por consiguiente W es un subespacio de V.

5.10. Sea V el espacio vectorial de los polinomios $a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ con coeficientes reales, es decir, $a_i \in \mathbf{R}$. Determinar si W es o no un subespacio de V, donde:

a) W consiste en todos los polinomios con coeficientes enteros.

b) W consiste en todos los polinomios de grado ≤ 3.

c) W consiste en todos los polinomios en los que sólo aparecen potencias pares de t.

d) W consiste en todos los polinomios que tienen por raíz a $\sqrt{-1}$.

a) No, porque los múltiplos escalares de los vectores en W no siempre pertenecen a W. Por ejemplo,

$$v = 3 + 5t + 7t^2 \in W \qquad \text{pero} \qquad \tfrac{1}{2}v = \tfrac{3}{2} + \tfrac{5}{2}t + \tfrac{7}{2}t^2 \notin W$$

(Obsérvese que W es «cerrado» bajo la suma vectorial, esto es, las sumas de elementos de W pertenecen a W.)

b), *c*) y *d*) Sí, puesto que, en cada caso, W es no vacío, la suma de elementos de W pertenece a W y los múltiplos escalares de cualquier elemento de W pertenecen a W.

5.11. Demostrar el Teorema 5.4.

Sean $\{W_i : i \in I\}$ una colección de subespacios de V y $W = \bigcap (W_i : i \in I)$. Dado que cada W_i es un subespacio, $0 \in W_i$ para todo $i \in I$. Por tanto, $0 \in W$. Supongamos que $u, v \in W$. En ese caso, $u, v \in W_i$ para todo $i \in I$. Como cada W_i es un subespacio, $au + bv \in W_i$ para cada $i \in I$. Por consiguiente, $au + bv \in W$. De este modo, W es un subespacio de V.

COMBINACIONES LINEALES. ENVOLVENTES LINEALES

5.12. Expresar $v = (1, -2, 5)$ en $\mathbf{R}^3$ como combinación lineal de los vectores u_1, u_2, u_3, siendo $u_1 = (1, -3, 2)$, $u_2 = (2, -4, -1)$, $u_3 = (1, -5, 7)$.

Primero tomamos

$$(1, -2, 5) = x(1, -3, 2) + y(2, -4, -1) + z(1, -5, 7) = (x + 2y + z, -3x - 4y - 5z, 2x - y + 7z)$$

Formamos el sistema de ecuaciones equivalente y lo reducimos a forma escalonada:

$$\begin{aligned} x + 2y + z &= 1 \\ -3x - 4y - 5z &= -2 \\ 2x - y + 7z &= 5 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + z &= 1 \\ 2y - 2z &= 1 \\ -5y + 5z &= 3 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + z &= 1 \\ 2y - 2z &= 1 \\ 0 &= 11 \end{aligned}$$

El sistema no tiene solución. De aquí que v no sea combinación lineal de u_1, u_2, u_3.

5.13. Expresar el polinomio sobre $\mathbf{R}$ $v = t^2 + 4t - 3$ como combinación lineal de los polinomios $p_1 = t^2 - 2t + 5$, $p_2 = 2t^2 - 3t$, $p_3 = t + 3$.

Tomamos v como combinación lineal de p_1, p_2, p_3 utilizando incógnitas x, y, z:

$$\begin{aligned} t^2 + 4t - 3 &= x(t^2 - 2t + 5) + y(2t^2 - 3t) + z(t + 3) = \\ &= xt^2 - 2xt + 5x + 2yt^2 - 3yt + zt + 3z = \\ &= (x + 2y)t^2 + (-2x - 3y + z)t + (5x + 3z) \end{aligned}$$

Igualamos entre sí los coeficientes de las mismas potencias de t y reducimos el sistema a forma escalonada:

$$\begin{aligned} x + 2y &= 1 \\ -2x - 3y + z &= 4 \\ 5x + 3z &= -3 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y &= 1 \\ y + z &= 6 \\ -10y + 3z &= -8 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y &= 1 \\ y + z &= 6 \\ 13z &= 52 \end{aligned}$$

El sistema está en forma triangular y tiene solución. Resolver por sustitución hacia atrás conduce a $x = -3$, $y = 2$, $z = 4$. Así pues, $v = -3p_1 + 2p_2 + 4p_3$.

5.14. Escribir la matriz $E = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ como combinación lineal de las matrices

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \qquad \text{y} \qquad C = \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix}$$

Tomamos E como combinación lineal de A, B, C usando las incógnitas x, y, z: $E = xA + yB + zC$.

$$\begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} = x\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + z\begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} =$$
$$= \begin{pmatrix} x & x \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ y & y \end{pmatrix} + \begin{pmatrix} 0 & 2z \\ 0 & -z \end{pmatrix} = \begin{pmatrix} x & x + 2z \\ x + y & y - z \end{pmatrix}$$

Construimos el sistema de ecuaciones equivalente igualando entre sí las entradas correspondientes:

$$x = 3 \qquad x + y = 1 \qquad x + 2z = 1 \qquad y - z = -1$$

Sustituimos $x = 3$ en las ecuaciones segunda y tercera para obtener $y = -2$ y $z = -1$. Como estos valores también satisfacen la última ecuación, forman una solución del sistema. Por tanto, $E = 3A - 2B - C$.

5.15. Encontrar una condición a imponer sobre a, b, c para que $w = (a, b, c)$ sea combinación lineal de $u = (1, -3, 2)$ y $v = (2, -1, 1)$, esto es, para que w pertenezca a $\text{lin}\,(u, v)$.

Tomamos $w = xu + yv$ utilizando incógnitas x e y:

$$(a, b, c) = x(1, -3, 2) + y(2, -1, 1) = (x + 2y, -3x - y, 2x + y)$$

Construimos el sistema equivalente y lo reducimos a forma escalonada:

$$\begin{aligned} x + 2y &= a \\ -3x - y &= b \\ 2x + y &= c \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y &= a \\ 5y &= 3a + b \\ -3y &= -2a + c \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y &= a \\ 5y &= 3a + b \\ 0 &= -a + 3b + 5c \end{aligned}$$

El sistema es compatible si y sólo si $a - 3b - 5c = 0$, luego w es combinacional lineal de u y v cuando $a - 3b - 5c = 0$.

5.16. Demostrar el Teorema 5.6.

Supongamos que S es vacío. Por definición, $\text{lin}\, S = \{0\}$. Por consiguiente, $\text{lin}\, S = \{0\}$ es un subespacio y $S \subseteq \text{lin}\, S$. Supongamos que S es no vacío y que $v \in S$. En ese caso, $1v = v \in \text{lin}\, S$; por tanto, S es un subconjunto de $\text{lin}\, S$. Siendo así, $\text{lin}\, S$ es no vacío porque S es no vacío. Supongamos ahora que $v, w \in \text{lin}\, S$, por ejemplo

$$v = a_1 v_1 + \cdots + a_m v_m \qquad \text{y} \qquad w = b_1 w_1 + \cdots + b_n w_n$$

donde $v_i, w_j \in S$ y a_i, b_j son escalares. Entonces

$$v + w = a_1 v_1 + \cdots + a_m v_m + b_1 w_1 + \cdots + b_n w_n$$

y, para cualquier escalar k,

$$kv = k(a_1 v_1 + \cdots + a_m v_m) = ka_1 v_1 + \cdots + ka_m v_m$$

pertenecen a $\text{lin}\, S$, puesto que cada uno es combinación lineal de vectores en S. Así $\text{lin}\, S$ es un subespacio de V.

Ahora supongamos que W es un subespacio de V que contiene S y que $v_1, \ldots, v_m \in S \subseteq W$. En tal caso, todos los múltiplos $a_1 v_1, \ldots, a_m v_m \in W$, donde $a_i \in K$, y la suma $a_1 v_1 + \cdots + a_m v_m \in W$. Esto es, W contiene todas las combinaciones lineales de elementos de S. En consecuencia, lin S es un subespacio de W, como se pretendía.

DEPENDENCIA LINEAL

5.17. Determinar si $u = 1 - 3t + 2t^2 - 3t^3$ y $v = -3 + 9t - 6t^2 + 9t^3$ son linealmente dependientes.

Dos vectores son linealmente dependientes si y sólo si uno es un múltiplo del otro. En este caso, $v = -3u$.

5.18. Determinar si los siguientes vectores en $\mathbf{R}^3$ son o no linealmente dependientes:

$$u = (1, -2, 1),\ v = (2, 1, -1),\ w = (7, -4, 1).$$

Método 1. Igualamos a cero una combinación lineal de los vectores con escalares desconocidos x, y y z:

$$x(1, -2, 1) + y(2, 1, -1) + z(7, -4, 1) = (0, 0, 0)$$

Entonces

$$(x, -2x, x) + (2y, y, -y) + (7z, -4z, z) = (0, 0, 0)$$

o sea,

$$(x + 2y + 7z, -2x + y - 4z, x - y + z) = (0, 0, 0)$$

Igualamos entre sí las componentes correspondientes para llegar al sistema homogéneo equivalente y lo reducimos a forma escalonada:

$$\begin{aligned} x + 2y + 7z &= 0 \\ -2x + y - 4z &= 0 \\ x - y + z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + 7z &= 0 \\ 5y + 10z &= 0 \\ -3y - 6z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + 2y + 7z &= 0 \\ y + 2z &= 0 \end{aligned}$$

El sistema, en forma escalonada, tiene únicamente dos ecuaciones no nulas en las tres incógnitas, de aquí que el sistema tenga una solución no nula. Así que los vectores originales son linealmente dependientes.

Método 2. Construimos la matriz cuyas filas son los vectores dados y la reducimos a forma escalonada utilizando las operaciones elementales entre filas:

$$\begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 7 & -4 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 10 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Dado que la matriz escalonada tiene una fila nula, los vectores son linealmente dependientes. (Los tres vectores dados generan un espacio de dimensión 2.)

5.19. Considérese el espacio vectorial $\mathbf{P}(t)$ de los polinomios sobre $\mathbf{R}$. Determinar si los polinomios u, v y w son linealmente dependientes, siendo $u = t^3 + 4t^2 - 2t + 3$, $v = t^3 + 6t^2 - t + 4$, $w = 3t^3 + 8t^2 - 8t + 7$.

Igualamos al polinomio cero una combinación lineal de u, v y w con escalares desconocidos x, y y z; esto es, tomamos $xu + yv + zw = 0$. Entonces

$$x(t^3 + 4t^2 - 2t + 3) + y(t^3 + 6t^2 - t + 4) + z(3t^3 + 8t^2 - 8t + 7) = 0$$

o
$$xt^3 + 4xt^2 - 2xt + 3x + yt^3 + 6yt^2 - yt + 4y + 3zt^3 + 8zt^2 - 8zt + 7z = 0$$

o
$$(x + y + 3z)t^3 + (4x + 6y + 8z)t^2 + (-2x - y - 8z)t + (3x + 4y + 7z) = 0$$

Igualamos a cero los coeficientes de cada potencia de t y reducimos el sistema a forma escalonada:

$$\begin{aligned} x + y + 3z &= 0 \\ 4x + 6y + 8z &= 0 \\ -2x - y - 8z &= 0 \\ 3x + 4y + 7z &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x + y + 3z &= 0 \\ 2y - 4z &= 0 \\ y - 2z &= 0 \\ y - 2z &= 0 \end{aligned} \quad \text{o finalmente} \quad \begin{aligned} x + y + 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

El sistema en forma escalonada tiene una variable libre y por tanto alguna solución no nula. Hemos probado que $xu + yv + zw = 0$ no implica $x = 0$, $y = 0$, $z = 0$, luego los polinomios son linealmente independientes.

5.20. Sea V el espacio vectorial de las funciones de $\mathbf{R}$ en $\mathbf{R}$. Probar que $f, g, h \in V$ son linealmente independientes, donde $f(t) = \operatorname{sen} t$, $g(t) = \cos t$, $h(t) = t$.

Igualamos a la función cero, 0, una combinación lineal de las funciones con escalares desconocidos x, y y z: $xf + yg + zh = 0$, para luego demostrar que, necesariamente, $x = 0$, $y = 0$, $z = 0$. Hacemos énfasis en que $xf + yg + xh = 0$ quiere decir que, *para todo valor de* t, $xf(t) + yg(t) + xh(t) = 0$.

En la ecuación $x \operatorname{sen} t + y \cos t + zt = 0$ sustituimos

$t = 0$	para obtener	$x \cdot 0 + y \cdot 1 + z \cdot 0 = 0$	o sea	$y = 0$
$t = \pi/2$	para obtener	$x \cdot 1 + y \cdot 0 + z\pi/2 = 0$	o sea	$x + \pi z/2 = 0$
$t = \pi$	para obtener	$x \cdot 0 + y(-1) + z \cdot \pi = 0$	o sea	$-y + \pi z = 0$

Resolvemos el sistema $\begin{cases} y = 0 \\ x + \pi z/2 = 0 \\ -y + \pi z = 0 \end{cases}$ y conseguimos la solución única: $x = 0$, $y = 0$, $z = 0$.

Por consiguiente, f, g y h son linealmente independientes.

5.21. Sea V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$. Determinar si las matrices A, B, $C \in V$ son linealmente dependientes, siendo:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Igualamos a la matriz cero una combinación lineal de A, B y C con escalares desconocidos x, y y z; es decir, tomamos $xA + yB + zC = 0$.

$$x\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + y\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + z\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

es decir,

$$\begin{pmatrix} x + y + z & x + z \\ x & x + y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Igualamos entre sí las componentes correspondientes para llegar al sistema de ecuaciones lineales equivalente:

$$x + y + z = 0 \qquad x + z = 0 \qquad x = 0 \qquad x + y = 0$$

Resolviendo el sistema obtenemos únicamente la solución nula: $x = 0$, $y = 0$, $z = 0$. Hemos demostrado que $xA + yB + zC = 0$ implica $x = 0$, $y = 0$, $z = 0$; por tanto, las matrices A, B y C son linealmente independientes.

5.22. Supóngase que u, v y w son vectores linealmente independientes. Probar que también lo son $u + v$, $u - v$ y $u - 2v + w$.

Supongamos $x(u + v) + y(u - v) + z(u - 2v + w) = 0$, donde x, y y z son escalares. Entonces

$$xu + xv + yu - yv + zu - 2zv + zw = 0$$

o sea,

$$(x + y + z)u + (x - y - 2z)v + zw = 0$$

Pero u, v y w son linealmente independientes, por lo que los coeficientes en la relación anterior son todos 0:

$$\begin{aligned} x + y + z &= 0 \\ x - y - 2z &= 0 \\ z &= 0 \end{aligned}$$

La única solución del sistema precedente es $x = 0$, $y = 0$, $z = 0$. En consecuencia, $u + v$, $u - v$ y $u - 2v + w$ son linealmente independientes.

5.23. Demostrar que los vectores $v = (1 + i,\ 2i)$ y $w = (1,\ 1 + i)$ en $\mathbf{C}^2$ son linealmente dependientes sobre el cuerpo complejo $\mathbf{C}$, pero linealmente independientes sobre el cuerpo real $\mathbf{R}$.

Recordemos que dos vectores son linealmente dependientes (sobre un cuerpo K) si y sólo si uno de ellos es múltiplo de otro (por un elemento de K). Como

$$(1 + i)w = (1 + i)(1,\ 1 + i) = (1 + i,\ 2i) = v$$

v y w son linealmente dependientes sobre $\mathbf{C}$. Por el contrario, v y w son linealmente independientes sobre $\mathbf{R}$, ya que ningún múltiplo real de w puede ser igual a v. Concretamente, cuando k es real, la primera componente de $kw = (k,\ k + ki)$ es real y nunca puede ser igual a la primera componente $1 + i$ de v, que es compleja.

BASES Y DIMENSION

5.24. Determinar si $(1, 1, 1)$, $(1, 2, 3)$ y $(2, -1, 1)$ constituyen una base del espacio vectorial $\mathbf{R}^3$.

Los tres vectores forman una base si y sólo si son linealmente independientes. Construimos la matriz A cuyas filas son los vectores dados y la reducimos por filas a forma escalonada:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -3 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{pmatrix}$$

La matriz escalonada no tiene filas nulas, luego los tres vectores son linealmente independientes y en consecuencia forman una base de $\mathbf{R}^3$.

5.25. Determinar si (1, 1, 1, 1), (1, 2, 3, 2), (2, 5, 6, 4) y (2, 6, 8, 5) forman una base de $\mathbf{R}^4$.

Construimos la matriz A cuyas filas son los vectores dados y la reducimos por filas a forma escalonada:

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 2 & 5 & 6 & 4 \\ 2 & 6 & 8 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 3 & 4 & 2 \\ 0 & 4 & 6 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & -2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

La matriz escalonada tiene una fila nula; por consiguiente, los cuatro vectores son linealmente dependientes y no constituyen una base de $\mathbf{R}^4$.

5.26. Considérese el espacio vectorial $\mathbf{P}_n(t)$ de los polinomios en t de grado $\leq n$. Determinar si $1 + t$, $t + t^2$, $t^2 + t^3$, ..., $t^{n-1} + t^n$ forman o no una base de $\mathbf{P}_n(t)$.

Los polinomios son linealmente independientes, pues cada uno de ellos es de grado mayor que los precedentes. Sin embargo, hay sólo n polinomios y dim $\mathbf{P}_n(t) = n + 1$. Por tanto, los polinomios no constituyen una base de $\mathbf{P}_n(t)$.

5.27. Sea V el espacio vectorial de las matrices reales 2×2. Determinar si

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

constituyen una base de V.

Los vectores coordenados (véase la Sección 5.10) de las matrices respecto a la base usual son, respectivamente,

$$[A] = (1, 1, 0, 0) \qquad [B] = (0, 1, 1, 0) \qquad [C] = (0, 0, 1, 1) \qquad [D] = (0, 0, 0, 1)$$

Estos vectores coordenados forman una matriz en forma escalonada y por consiguiente son linealmente independientes. De este modo, las cuatro matrices correspondientes son linealmente independientes. Más aún, como dim $V = 4$, forman una base de V.

5.28. Sea V el espacio vectorial de las matrices simétricas 2×2 sobre K. Mostrar que dim $V = 3$. [Recuérdese que $A = (a_{ij})$ es simétrica si y sólo si $A = A^T$ o, equivalentemente, $a_{ij} = a_{ji}$.]

Una matriz simétrica 2×2 arbitraria es de la forma $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, donde $a, b, c \in K$. (Nótese que hay tres «variables».) Tomando

i) $a = 1, b = 0, c = 0$, ii) $a = 0, b = 1, c = 0$, iii) $a = 0, b = 0, c = 1$

obtenemos las respectivas matrices

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Demostremos que $\{E_1, E_2, E_3\}$ es una base de V, o sea, que a) genera V y b) es linealmente independiente.

a) Para la matriz arbitraria anterior A en V tenemos

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = aE_1 + bE_2 + cE_3$$

Así pues, $\{E_1, E_2, E_3\}$ genera V.

b) Supongamos que $xE_1 + yE_2 + zE_3 = 0$, donde x, y, z son escalares desconocidos. Esto es, supongamos que

$$x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + z\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{o sea} \quad \begin{pmatrix} x & y \\ y & z \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Igualando entre sí las componentes correspondientes obtenemos $x = 0$, $y = 0$, $z = 0$. En otras palabras,

$$xE_1 + yE_2 + zE_3 = 0 \qquad \text{implica} \qquad x = 0,\ y = 0,\ z = 0$$

De acuerdo con ello, $\{E_1, E_2, E_3\}$ es linealmente independiente.

De este modo, $\{E_1, E_2, E_3\}$ es una base de V y por tanto la dimensión de V es 3.

5.29. Considérese el cuerpo complejo **C** que contiene el cuerpo real **R** que, a su vez, contiene el cuerpo racional **Q**. (**C** es un espacio vectorial sobre **R** y **R** un espacio vectorial sobre **Q**.)

a) Probar que **C** es un espacio de dimensión 2 sobre **R**.

b) Probar que **R** es un espacio de dimensión infinita sobre **Q**.

a) Afirmamos que $\{1, i\}$ es una base de **C** sobre **R**. Ello se debe a que si $v \in \mathbf{C}$, $v = a + bi = a \cdot 1 + b \cdot i$ con a, $b \in \mathbf{R}$; es decir, $\{1, i\}$ genera **C** sobre **R**. Además, si $x \cdot 1 + y \cdot i = 0$ o $x + yi = 0$, donde x, $y \in \mathbf{R}$, necesariamente $x = 0$ e $y = 0$; esto es, $\{1, i\}$ es linealmente independiente sobre **R**. De este modo, $\{1, i\}$ es base de **C** sobre **R** y **C** es de dimensión 2 sobre **R**.

b) Afirmamos que, para todo n, $\{1, \pi, \pi^2, \ldots, \pi^n\}$ es linealmente independiente sobre **Q**. Supongamos $a_0 1 + a_1\pi + a_2\pi^2 + \cdots + a_n\pi^n = 0$, donde los $a_i \in \mathbf{Q}$ y no todos los a_i son 0. En tal caso, π es una raíz del siguiente polinomio no nulo sobre **Q**: $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Pero puede demostrarse que π es un número trascendente, o sea, que no es raíz de ningún polinomio no nulo sobre **Q**. En consecuencia, los $n + 1$ números reales $1, \pi, \pi^2, \ldots, \pi^n$ son linealmente independientes sobre **Q**. Así, para cualquier n finito, **R** no puede ser de dimensión n sobre **Q**; **R** es de dimensión infinita sobre **Q**.

5.30. Sea $S = \{u_1, u_2, \ldots, u_n\}$ un subconjunto de un espacio vectorial V. Demostrar que las dos condiciones siguientes son equivalentes: *a*) S es linealmente independiente y genera V, *b*) todo vector $v \in V$ puede escribirse de forma única como combinación lineal de vectores de S.

Supongamos que se verifica *a*). Como S genera V, el vector v será combinación lineal de los u_i; digamos

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

Supongamos que tenemos también

$$v = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

Restando llegamos a

$$0 = v - v = (a_1 - b_1)u_1 + (a_2 - b_2)u_2 + \cdots + (a_n - b_n)u_n$$

Ahora bien, los u_i son linealmente independientes; por tanto, los coeficientes en la relación anterior son todos 0:

$$a_1 - b_1 = 0, a_2 - b_2 = 0, \ldots, a_n - b_n = 0$$

Por consiguiente, $a_1 = b_1$, $a_2 = b_2$, ..., $a_n = b_n$; luego la representación de v como combinación lineal de los u_i es única. Siendo así, *a*) implica *b*).

Supongamos que se verifica *b*). Entonces S genera V. Supongamos

$$0 = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

No obstante, tenemos

$$0 = 0u_1 + 0u_2 + \cdots + 0u_n$$

Por hipótesis, la representación de 0 como combinación lineal de los u_i es única, por lo que cada $c_i = 0$ y los u_i son linealmente independientes. Así *b*) implica *a*).

DIMENSION Y SUBESPACIOS

5.31. Hallar una base y la dimensión del subespacio W de $\mathbf{R}^3$, donde:

a) $W = \{(a, b, c): a + b + c = 0\}$, *b*) $W = \{(a, b, c): a = b = c\}$,

c) $W = (a, b, c): c = 3a\}$

a) Nótese que $W \neq \mathbf{R}^3$, dado que, por ejemplo, $(1, 2, 3) \notin W$. Por tanto, $\dim W < 3$. Nótese, asimismo, que $u_1 = (1, 0, -1)$ y $u_2 = (0, 1, -1)$ son dos vectores linealmente independientes en W. De este modo, $\dim W = 2$ y u_1 y u_2 forman una base de W.

b) El vector $u = (1, 1, 1) \in W$. Todo vector $w \in W$ es de la forma $w = (k, k, k)$, luego $w = ku$. Así u genera W y $\dim W = 1$.

c) $W \neq \mathbf{R}^3$ porque, por ejemplo, $(1, 1, 1) \notin W$. Entonces $\dim W < 3$. Los vectores $u_1 = (1, 0, 3)$ y $u_2 = (0, 1, 0)$ pertenecen a W y son linealmente independientes. Por tanto, $\dim W = 2$ y u_1 y u_2 constituyen una base de W.

5.32. Encontrar una base y la dimensión del subespacio W de $\mathbf{R}^4$ generado por

$$u_1 = (1, -4, -2, 1), \quad u_2 = (1, -3, -1, 2), \quad u_3 = (3, -8, -2, 7)$$

Aplicamos el Algoritmo *5.8A* del espacio fila. Construimos una matriz en la que las filas son los vectores dados y la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & -4 & -2 & 1 \\ 1 & -3 & -1 & 2 \\ 3 & -8 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 4 & 4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -4 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Las filas no nulas de la matriz escalonada forman una base de W, de modo que $\dim W = 2$. En particular, esto significa que los tres vectores originales son linealmente dependientes.

5.33. Sea W el subespacio de $\mathbf{R}^4$ generado por los vectores

$$u_1 = (1, -2, 5, -3), \quad u_2 = (2, 3, 1, -4) \quad u_3 = (3, 8, -3, -5)$$

a) Hallar una base y la dimensión de W. *b*) Extender la base de W a una del espacio $\mathbf{R}^4$ completo.

a) Construimos la matriz A cuyas filas son los vectores dados y la reducimos por filas a forma escalonada:

$$A = \begin{pmatrix} 1 & -2 & 5 & -3 \\ 2 & 3 & 1 & -4 \\ 3 & 8 & -3 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 14 & -18 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Las filas no nulas $(1, -2, 5, -3)$ y $(0, 7, -9, 2)$ de la matriz escalonada forman una base del espacio fila de A y por ende W. Así, en particular, $\dim W = 2$.

b) Buscamos cuatro vectores linealmente independientes que incluyan los dos anteriores. Los vectores $(1, -2, 5, -3)$, $(0, 7, -9, 2)$, $(0, 0, 1, 0)$ y $(0, 0, 0, 1)$ son linealmente independientes (porque forman una matriz escalonada) y constituyen, pues, una base de $\mathbf{R}^4$ que es una extensión de la de W.

5.34. Sea W el subespacio de $\mathbf{R}^5$ generado por los vectores $u_1 = (1, 2, -1, 3, 4)$, $u_2 = (2, 4, -2, 6, 8)$, $u_3 = (1, 3, 2, 2, 6)$, $u_4 = (1, 4, 5, 1, 8)$ y $u_5 = (2, 7, 3, 3, 9)$. Hallar un subconjunto de los vectores que sea base de W.

Método 1. Aquí empleamos el Algoritmo 5.8*B* de expulsión. Construimos la matriz cuyas columnas son los vectores dados y la reducimos a forma escalonada:

$$\begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 2 & 4 & 3 & 4 & 7 \\ -1 & -2 & 2 & 5 & 3 \\ 3 & 6 & 2 & 1 & 3 \\ 4 & 8 & 6 & 8 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 6 & 5 \\ 0 & 0 & -1 & -2 & -3 \\ 0 & 0 & 2 & 4 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Las posiciones de los pivotes están en las columnas C_1, C_3 y C_5. Por consiguiente, los vectores correspondientes u_1, u_3 y u_5 forman una base de W y $\dim W = 3$.

Método 2. Aquí utilizamos una ligera modificación del Algoritmo 5.8*A* del espacio fila. Construimos la matriz cuyas filas son los vectores dados y la reducimos a forma «escalonada» pero sin intercambiar ninguna fila nula:

$$\begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 2 & 4 & -2 & 6 & 8 \\ 1 & 3 & 2 & 2 & 6 \\ 1 & 4 & 5 & 1 & 8 \\ 2 & 7 & 3 & 3 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -1 & 2 \\ 0 & 2 & 6 & -2 & 4 \\ 0 & 3 & 5 & -3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 & -5 \end{pmatrix}$$

Las filas no nulas son la primera, tercera y quinta; luego u_1, u_3 y u_5 forman una base de W. De este modo, en particular, $\dim W = 3$.

5.35. Sea V el espacio vectorial de las matrices reales 2×2. Encontrar la dimensión y una base del subespacio W de V generado por

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 5 \\ 1 & -1 \end{pmatrix} \qquad C = \begin{pmatrix} 5 & 12 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 3 & 4 \\ -2 & 5 \end{pmatrix}$$

Los vectores coordenados (véase la Sección 5.10) de las matrices dadas respecto a la base usual de V son los siguientes:

$$[A] = [1, 2, -1, 3] \qquad [B] = [2, 5, 1, -1] \qquad [C] = [5, 12, 1, 1] \qquad [D] = [3, 4, -2, 5]$$

Formamos una matriz cuyas filas sean los vectores coordenados y la reducimos a forma escalonada:

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 5 & 1 & -1 \\ 5 & 12 & 1 & 1 \\ 3 & 4 & -2 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 2 & 6 & -14 \\ 0 & -2 & 1 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & -18 \end{pmatrix}$$

Las filas no nulas son linealmente independientes, por lo que las matrices asociadas $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 3 & -7 \end{pmatrix}$ y $\begin{pmatrix} 0 & 0 \\ 7 & -18 \end{pmatrix}$ constituyen una base de W y dim $W = 3$. (Nótese también que A, B y D forman una base de W.)

TEOREMAS SOBRE DEPENDENCIA LINEAL, BASES Y DIMENSION

5.36. Demostrar el Lema 5.10.

Como los v_i son linealmente dependientes, existen escalares $a_1, \ldots, a_m$, no todos 0, tales que $a_1 v_1 + \cdots + a_m v_m = 0$. Sea k el mayor entero tal que $a_k \neq 0$. En ese caso,

$$a_1 v_1 + \cdots + a_k v_k + 0v_{k+1} + \cdots + 0v_m = 0 \qquad \text{o} \qquad a_1 v_1 + \cdots + a_k v_k = 0$$

Supongamos que $k = 1$; entonces $a_1 v_1 = 0$, $a_1 \neq 0$, luego $v_1 = 0$. Pero los v_i son vectores no nulos; por tanto, $k > 1$ y

$$v_k = -a_k^{-1} a_1 v_1 - \cdots - a_k^{-1} a_{k-1} v_{k-1}$$

Esto es, v_k es una combinación lineal de los vectores precedentes.

5.37. Demostrar el Teorema 5.11.

Supongamos que $R_n, R_{n-1}, \ldots, R_1$ son linealmente dependientes. Una de las filas, digamos R_m, debe ser combinación lineal de las precedentes:

$$R_m = a_{m+1} R_{m+1} + a_{m+2} R_{m+2} + \cdots + a_n R_n$$

Sea ahora la k-ésima componente de R_m su primera entrada no nula. En tal caso, dado que la matriz está en forma escalonada, las componentes k-ésimas de $R_{m+1}, \ldots, R_n$ son todas 0 y la componente k-ésima de (*) es, pues,

$$a_{m+1} \cdot 0 + a_{m+2} \cdot 0 + \cdots + a_n \cdot 0 = 0$$

Pero esto contradice la suposición de que la componente k-ésima de R_m no es 0. Por consiguiente, $R_1, \ldots, R_m$ son linealmente independientes.

5.38. Supóngase que $\{v_1, \ldots, v_m\}$ genera un espacio vectorial V. Se pide demostrar:

a) Si $w \in V$, entonces $\{w, v_1, \ldots, v_m\}$ es linealmente dependiente y genera V.

b) Si v_i es una combinación lineal de los vectores $(v_1, v_2, \ldots, v_{i-1})$, entonces $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ genera V.

a) El vector w es una combinación lineal de los v_i ya que $\{v_i\}$ genera V. De acuerdo con ello, $\{w, v_1, \ldots, v_m\}$ es linealmente dependiente. Obviamente, w y los v_i generan V, puesto que ya lo hacían los v_i por sí mismos. O sea, $\{w, v_1, \ldots, v_m\}$ genera V.

b) Supongamos $v_i = k_1 v_1 + \cdots + k_{i-1} v_{i-1}$. Sea $u \in V$. Como $\{v_i\}$ genera V, u es una combinación lineal de los v_i, por ejemplo, $u = a_1 v_1 + \cdots + a_m v_m$. Sustituyendo v_i obtenemos

$$\begin{aligned} u &= a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_i(k_1 v_1 + \cdots + k_{i-1} v_{i-1}) + a_{i+1} v_{i+1} + \cdots + a_m v_m = \\ &= (a_1 + a_i k_1) v_1 + \cdots + (a_{i-1} + a_i k_{i-1}) v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_m v_m \end{aligned}$$

Así $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ genera V. En otras palabras, podemos suprimir v_i del conjunto generador y tener todavía un conjunto generador.

5.39. Demostrar el Lema 5.13.

Basta demostrarlo en el caso de que los v_i no sean todos 0. (Pruébese.) Como los v_i generan V, tenemos, de acuerdo con el Problema 5.38, que

$$\{w_1, v_1, \ldots, v_n\} \qquad [1]$$

es linealmente dependiente y genera V. Por el Lema 5.10 sabemos que uno de los vectores en [1] es una combinación lineal de los precedentes. Este vector no puede ser w_1, luego debe ser uno de los v, digamos v_j. De este modo, según el Problema 5.38, podemos suprimir v_j del conjunto generador [1] obteniendo el conjunto generador

$$\{w_1; v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\} \qquad [2]$$

Ahora repetimos el argumento con el vector w_2. Es decir, como [2] genera V, el conjunto

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\} \qquad [3]$$

es linealmente dependiente y también genera V. De nuevo por el Lema 5.10 uno de los vectores es combinación lineal de los precedentes. Subrayamos el hecho de que este vector no puede ser w_1 ni w_2, porque $\{w_1, \ldots, w_m\}$ es independiente; por consiguiente, debe ser uno de los v, por ejemplo, v_k. En ese caso, según el problema precedente, podemos suprimir v_k del conjunto generador [3] obteniendo

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\}$$

Repetimos el argumento con w_3, y así sucesivamente. En cada paso somos capaces de añadir uno de los w y suprimir uno de los v en el conjunto generador. Si $m \leq n$, conseguiremos finalmente un conjunto generador de la forma requerida:

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Por último, probemos que $m > n$ no es posible. En caso contrario, tras n de los pasos anteriores, obtendríamos el conjunto generador $\{w_1, \ldots, w_n\}$. Esto implicaría que w_{n+1} es combinación lineal de $w_1, \ldots, w_n$, lo que contradice la hipótesis de que $\{w_i\}$ es linealmente independiente.

5.40. Demostrar el Teorema 5.12.

Supongamos que $\{u_1, u_2, \ldots, u_n\}$ y $\{v_1, v_2, \ldots\}$ son dos bases de V. Dado que $\{u_i\}$ genera V, la base $\{v_1, v_2, \ldots\}$ debe contener n o menos vectores, o bien ser linealmente dependiente, de acuerdo

con el Problema 5.39 (Lema 5.13). Por otra parte, si la base $\{v_1, v_2, \ldots\}$ tiene menos de n elementos, $\{u_1, u_2, \ldots, u_n\}$ es linealmente dependiente, según el Problema 5.39. Siendo así, la base $\{v_1, v_2, \ldots\}$ contiene exactamente n vectores y el teorema es, pues, cierto.

5.41. Demostrar el Teorema 5.14.

Supongamos que $B = \{w_1, w_2, \ldots, w_n\}$ es una base de V.

i) Como B genera V, $n + 1$ o más vectores arbitrarios son linealmente dependientes por el Lema 5.13.

ii) Según el Lema 5.13, pueden añadirse elementos de B a S para formar un conjunto generador de V con n elementos. Dado que S ya tiene n elementos, S es, en sí mismo, un conjunto generador de V. Por tanto, S es una base de V.

iii) Supongamos que T es linealmente dependiente. En ese caso, algún v_i es combinación lineal de los vectores precedentes. Por el Problema 5.38, V estará generado por los vectores de T sin v_i, que son $n - 1$. De acuerdo con el Lema 5.13, el conjunto independiente B no puede tener más de $n - 1$ elementos. Esto contradice el hecho de que B tiene n elementos, luego T es linealmente independiente y por ende es una base de V.

5.42. Demostrar el Teorema 5.15.

i) Supongamos que $\{v_1, \ldots, v_m\}$ es un subconjunto linealmente independiente maximal de S y que $w \in S$. En consecuencia, $\{v_1, \ldots, v_m, w\}$ es linealmente dependiente. Ningún v_k puede ser combinación lineal de los vectores precedentes, luego w es combinación lineal de los v_i. De este modo, $w \in \text{lin } v_i$ y por consiguiente $S \subseteq \text{lin } v_i$. Esto conduce a

$$V = \text{lin } S \subseteq \text{lin } v_i \subseteq V$$

Así pues, $\{v_i\}$ genera V y, como es linealmente independiente, es una base de V.

ii) Los restantes vectores forman un subconjunto linealmente independiente maximal de S y por la parte i) es una base de V.

5.43. Demostrar el Teorema 5.16.

Supongamos que $B = \{w_1, w_2, \ldots, w_n\}$ es una base de V. Entonces B genera V y por tanto V está generado por

$$S \cup B = \{u_1, u_2, \ldots, u_r, w_1, w_2, \ldots, w_n\}$$

De acuerdo con el Teorema 5.15, podemos suprimir de $S \cup B$ todo vector que sea combinación lineal de los precedentes para obtener una base B' de V. Como S es linealmente independiente, ningún u_k es combinación lineal de los vectores precedentes, por lo que B' contiene todos los vectores de S. Siendo así, S es parte de la base B' de V.

5.44. Demostrar el Teorema 5.17.

Dado que V es de dimensión n, $n + 1$ o más vectores son linealmente dependientes. Además, como toda base de W consiste en vectores linealmente independientes, no puede contener más de n elementos. De acuerdo con esto, $\dim W \leq n$.

En particular, si $\{w_1, \ldots, w_n\}$ es una base de W, por ser un conjunto independiente con n elementos, es también base de V. De este modo, $W = V$ cuando $\dim W = n$.

ESPACIO FILA Y RANGO DE UNA MATRIZ

5.45. Determinar si las siguientes matrices tienen el mismo espacio fila:

$$A=\begin{pmatrix}1 & 1 & 5\\ 2 & 3 & 13\end{pmatrix} \qquad B=\begin{pmatrix}1 & -1 & -2\\ 3 & -2 & -3\end{pmatrix} \qquad C=\begin{pmatrix}1 & -1 & -1\\ 4 & -3 & -1\\ 3 & -1 & 3\end{pmatrix}$$

Dos matrices tienen el mismo espacio fila si y sólo si sus formas canónicas por filas tienen las mismas filas no nulas; por consiguiente, reducimos cada matriz a forma canónica por filas:

$$A=\begin{pmatrix}1 & 1 & 5\\ 2 & 3 & 13\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 5\\ 0 & 1 & 3\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 2\\ 0 & 1 & 3\end{pmatrix}$$

$$B=\begin{pmatrix}1 & -1 & -2\\ 3 & -2 & -3\end{pmatrix}\sim\begin{pmatrix}1 & -1 & -2\\ 0 & 1 & 3\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 3\end{pmatrix}$$

$$C=\begin{pmatrix}1 & -1 & -1\\ 4 & -3 & -1\\ 3 & -1 & 3\end{pmatrix}\sim\begin{pmatrix}1 & -1 & -1\\ 0 & 1 & 3\\ 0 & 2 & 6\end{pmatrix}\sim\begin{pmatrix}1 & -1 & -1\\ 0 & 1 & 3\\ 0 & 0 & 0\end{pmatrix} \quad \text{a} \quad \begin{pmatrix}1 & 0 & 2\\ 0 & 1 & 3\\ 0 & 0 & 0\end{pmatrix}$$

Dado que las filas no nulas de las formas reducidas de A y C son las mismas, A y C tienen el mismo espacio fila. Por el contrario, las filas no nulas de la forma reducida de B no coinciden con las de las otras, por lo que B tiene un espacio fila diferente.

5.46. Probar que $A=\begin{pmatrix}1 & 3 & 5\\ 1 & 4 & 3\\ 1 & 1 & 9\end{pmatrix}$ y $B=\begin{pmatrix}1 & 2 & 3\\ -2 & -3 & -4\\ 7 & 12 & 17\end{pmatrix}$ tienen el mismo espacio columna.

Observemos que A y B tienen el mismo espacio columna si y sólo si las traspuestas A^T y B^T tienen el mismo espacio fila. Siendo así, reducimos A^T y B^T a forma canónica por filas:

$$A^T=\begin{pmatrix}1 & 1 & 1\\ 3 & 4 & 1\\ 5 & 3 & 9\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 1\\ 0 & 1 & -2\\ 0 & -2 & 4\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 1\\ 0 & 1 & -2\\ 0 & 0 & 0\end{pmatrix} \quad \text{a} \quad \begin{pmatrix}1 & 0 & 3\\ 0 & 1 & -2\\ 0 & 0 & 0\end{pmatrix}$$

$$B^T=\begin{pmatrix}1 & -2 & 7\\ 2 & -3 & 12\\ 3 & -4 & 17\end{pmatrix}\sim\begin{pmatrix}1 & -2 & 7\\ 0 & 1 & -2\\ 0 & 2 & -4\end{pmatrix}\sim\begin{pmatrix}1 & -2 & 7\\ 0 & 1 & -2\\ 0 & 0 & 0\end{pmatrix} \quad \text{a} \quad \begin{pmatrix}1 & 0 & 3\\ 0 & 1 & -2\\ 0 & 0 & 0\end{pmatrix}$$

Como A^T y B^T tienen el mismo espacio fila, A y B tienen el mismo espacio columna.

5.47. Considérense los subespacios $U = \text{lin}\,(u_1, u_2, u_3)$ y $W = \text{lin}\,(w_1, w_2, w_3)$ de $\mathbf{R}^3$, donde:

$$u_1 = (1, 1, -1), \qquad u_2 = (2, 3, -1), \qquad u_3 = (3, 1, -5)$$
$$w_1 = (1, -1, -3), \qquad w_2 = (3, -2, -8), \qquad w_3 = (2, 1, -3)$$

Probar que $U = W$.

Construimos la matriz A cuyas filas son los u_i y la reducimos a forma canónica por filas:

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & -1 \\ 3 & 1 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

A continuación construimos la matriz B cuyas filas son los w_i y la reducimos a forma canónica por filas:

$$B = \begin{pmatrix} 1 & -1 & -3 \\ 3 & -2 & -8 \\ 2 & 1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 3 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Dado que A y B tienen la misma forma canónica por filas, sus espacios fila son iguales y en consecuencia $U = W$.

5.48. Hallar el rango de la matriz A, donde:

$$a)\quad A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{pmatrix}, \qquad b)\quad A = \begin{pmatrix} 1 & 3 \\ 0 & -2 \\ 5 & -1 \\ -2 & 3 \end{pmatrix}$$

a) Como los rangos por filas y columnas son iguales, es más fácil tomar la traspuesta de A y después reducirla por filas a forma escalonada:

$$A^T = \begin{pmatrix} 1 & 2 & -2 & -1 \\ 2 & 1 & -1 & 4 \\ -3 & 0 & 3 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 6 & -3 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 0 & 3 & 7 \end{pmatrix}$$

Así rango $A = 3$.

b) Las dos columnas son linealmente independientes porque no son una múltiplo de la otra. Por tanto, rango $A = 2$.

5.49. Considérese una matriz arbitraria $A = (a_{ij})$. Supóngase que $u = (b_1, \ldots, b_n)$ es una combinación lineal de las filas $R_1, \ldots, R_m$ de A; por ejemplo, $u = k_1 R_1 + \cdots + k_m R_m$. Demostrar que

$$b_i = k_1 a_{1i} + k_2 a_{2i} + \cdots + k_m a_{mi} \qquad (i = 1, 2, \ldots, n)$$

donde $a_{1i}, \ldots, a_{mi}$ son las entradas de la i-ésima columna de A.

Se nos da $u = k_1 R_1 + \cdots + k_m R_m$; por tanto,

$$(b_1, \ldots, b_n) = k_1(a_{11}, \ldots, a_{1n}) + \cdots + k_m(a_{m1}, \ldots, a_{mn}) =$$
$$= (k_1 a_{11} + \cdots + k_m a_{m1}, \ldots, k_1 a_{m1} + \cdots + k_m a_{mn})$$

Igualando entre sí las componentes correspondientes, obtenemos el resultado deseado.

5.50. Supóngase que $A = (a_{ij})$ y $B = (b_{ij})$ son matrices escalonadas con entradas pivote:

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r} \quad \text{y} \quad b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$$

$$A = \begin{pmatrix} & a_{1j_1} & * & * & * & * & * & * \\ & & & a_{2j_2} & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & a_{rj_r} & * & * \\ & & & & & & & \end{pmatrix}, \qquad B = \begin{pmatrix} & b_{1k_1} & * & * & * & * & * & * & * \\ & & b_{2k_2} & * & * & * & * & * & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & & b_{sk_s} & * & * \\ & & & & & & & & \end{pmatrix}$$

Supóngase, asimismo, que A y B tienen el mismo espacio fila. Demostrar que las entradas pivote de A y B están en las mismas posiciones, esto es, demostrar que $j_1 = k_1$, $j_2 = k_2, \ldots, j_r = k_r$ y $r = s$.

Claramente, $A = 0$ si y sólo si $B = 0$, de manera que sólo es necesario probar el teorema cuando $r \geq 1$ y $s \geq 1$. Mostramos primero que $j_1 = k_1$. Supongamos que $j_1 < k_1$. En tal caso, la columna j_1-ésima de B es cero. Como la primera fila de A está en el espacio fila de B, tenemos, por el problema precedente

$$a_{1j_1} = c_1 0 + c_2 0 + \cdots + c_m 0 = 0$$

para ciertos escalares c_i. Pero esto contradice el hecho de que el elemento pivote $a_{1j_1} \neq 0$. Por consiguiente, $j_1 \geq k_1$ y, de forma similar, $k_1 \geq j_1$. Así pues, $j_1 = k_1$.

Sean ahora A' la submatriz de A obtenida suprimiendo su primera fila y B' la obtenida de B mediante el mismo procedimiento. Probamos que los espacios fila de A' y B' coinciden. El teorema se deducirá entonces por inducción, puesto que A' y B' son también matrices escalonadas.

Sean $R = (a_1, a_2, \ldots, a_n)$ cualquier fila de A' y $R_1, \ldots, R_m$ las filas de B. Dado que R está en el espacio fila de B, existen escalares $d_1, \ldots, d_m$ tales que $R = d_1 R_1 + d_2 R_2 + \cdots + d_m R_m$. Puesto que A está en forma escalonada y R no es su primera fila, la j_1-ésima entrada de R es cero: $a_i = 0$ para $i = j_1 = k_1$. Además, como B está en forma escalonada, todas las entradas en su k_1-ésima columna son 0 salvo la primera: $b_{1k_1} \neq 0$, pero $b_{2k_1} = 0, \ldots, b_{mk_1} = 0$. Siendo así,

$$0 = a_{k_1} = d_1 b_{1k_1} + d_2 0 + \cdots + d_m 0 = d_1 b_{1k_1}$$

Ahora $b_{1k_1} \neq 0$ y entonces $d_1 = 0$. De este modo, R es una combinación lineal de $R_2, \ldots, R_m$ y por tanto está en el espacio fila de B'. Como R era cualquier fila de A', el espacio fila de A' está contenido en el de B'. Similarmente, el espacio fila de B' está contenido en el de A'. Así A' y B' tienen el mismo espacio fila y queda demostrado el teorema.

5.51. Demostrar el Teorema 5.8.

Obviamente, si A y B tienen las mismas filas no nulas, necesariamente tienen el mismo espacio fila, de modo que sólo tenemos que demostrar el recíproco.

Supongamos que A y B tienen el mismo espacio fila y que $R \neq 0$ es la i-ésima fila de A. En ese caso existen escalares $c_1, \ldots, c_s$ tales que

$$R = c_1 R_1 + c_2 R_2 + \cdots + c_s R_s \qquad [1]$$

donde las R_i son las filas no nulas de B. El teorema quedará demostrado si probamos que $R = R_i$, o $c_i = 1$ pero $c_k = 0$ para $k \neq i$.

Sea a_{ij_i} la entrada pivote de R, o sea, la primera entrada no nula de R. Según [1] y el Problema 5.49,

$$a_{ij_i} = c_1 b_{1j_i} + c_2 b_{2j_i} + \cdots + c_s b_{sj_i} \qquad [2]$$

Pero por el problema precedente, b_{ij_i} es una entrada pivote de B y, como B está reducida por filas, es la única entrada distinta de cero en su columna j_i-ésima. Así pues, de [2] obtenemos $a_{ij_i} = c_i b_{ij_i}$. De cualquier modo, $a_{ij_i} = 1$ y $b_{ij_i} = 1$ porque A y B están reducidas por filas; por consiguiente, $c_i = 1$.

Supongamos que $k \neq i$ y que b_{kj_k} es la entrada pivote de R_k. De acuerdo con [1] y con el Problema 5.49,

$$a_{ij_k} = c_1 b_{1j_k} + c_2 b_{2j_k} + \cdots + c_s b_{sj_k} \tag{3}$$

Puesto que B está reducida por filas, b_{kj_k} es la única entrada distinta de cero en la columna j_k-ésima de B; por tanto, según [3], $a_{ij_k} = c_k b_{kj_k}$. Además, por el problema precedente, a_{kj_k} es una entrada pivote de A y, como A está reducida por filas, $a_{ij_k} = 0$. Siendo así, $c_k b_{kj_k} = 0$ y, dado que $b_{kj_k} = 1$, $c_k = 0$. En consecuencia, $R = R_i$ y el teorema queda demostrado.

5.52. Demostrar el Teorema 5.9.

Supongamos que A es equivalente por filas a dos matrices A_1 y A_2 en forma canónica por filas. Entonces f-lin A = f-lin A_1 y f-lin A = f-lin A_2. Por consiguiente, f-lin A_1 = f-lin A_2. Como A_1 y A_2 están en forma canónica por filas, $A_1 = A_2$ según el Teorema 5.8. Queda así demostrado el teorema.

5.53. Demostrar el Teorema 5.18.

Sea A una matriz $m \times n$ arbitraria:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Denotemos sus filas por $R_1, R_2, \ldots, R_m$:

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

Supongamos que el rango por filas es r y que los r vectores que siguen forman una base del espacio fila:

$$S_1 = (b_{11}, b_{12}, \ldots, b_{1n}), S_2 = (b_{21}, b_{22}, \ldots, b_{2n}), \ldots, S_r = (b_{r1}, b_{r2}, \ldots, b_{rn})$$

En tal caso, cada uno de los vectores fila es una combinación lineal de los S_i:

$$\begin{aligned} R_1 &= k_{11}S_1 + k_{12}S_2 + \cdots + k_{1r}S_r \\ R_2 &= k_{21}S_1 + k_{22}S_2 + \cdots + k_{2r}S_r \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ R_m &= k_{m1}S_1 + k_{m2}S_2 + \cdots + k_{mr}S_r \end{aligned}$$

donde los k_{ij} son escalares. Igualando entre sí las componentes i-ésimas en cada una de las ecuaciones vectoriales anteriores llegamos al siguiente sistema de ecuaciones, cada una válida para $i = 1, \ldots, n$:

$$\begin{aligned} a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots + k_{1r}b_{ri} \\ a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots + k_{2r}b_{ri} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots + k_{mr}b_{ri} \end{aligned}$$

Así, para $i = 1, \dots, n$,

$$\begin{pmatrix} a_{1i} \\ a_{2i} \\ \dots \\ a_{mi} \end{pmatrix} = b_{1i}\begin{pmatrix} k_{11} \\ k_{21} \\ \dots \\ k_{m1} \end{pmatrix} + b_{2i}\begin{pmatrix} k_{12} \\ k_{22} \\ \dots \\ k_{m2} \end{pmatrix} + \cdots + b_{ri}\begin{pmatrix} k_{1r} \\ k_{2r} \\ \dots \\ k_{mr} \end{pmatrix}$$

En otras palabras, cada una de las columnas de A es combinación lineal de los r vectores

$$\begin{pmatrix} k_{11} \\ k_{21} \\ \dots \\ k_{m1} \end{pmatrix}, \begin{pmatrix} k_{12} \\ k_{22} \\ \dots \\ k_{m1} \end{pmatrix}, \dots, \begin{pmatrix} k_{1r} \\ k_{2r} \\ \dots \\ k_{mr} \end{pmatrix}$$

De este modo, la dimensión del espacio columna de A es a lo sumo r, esto es, el rango por columnas es menor o igual que r y por ende menor o igual que el rango por filas.

De forma similar (o considerando la matriz traspuesta A^T) obtenemos que el rango por filas es menor o igual que el rango por columnas. Siendo así, los rangos por filas y columnas son iguales.

5.54. Supóngase que R es un vector fila y que A y B son matrices tales que RA y RB están definidos. Demostrar:

a) RB es una combinación lineal de las filas de B.

b) El espacio fila de AB está contenido en el de B.

c) El espacio columna de AB está contenido en el de A.

d) rango $AB \leq$ rango B y rango $AB \leq$ rango A.

a) Supongamos $R = (a_1, a_2, \dots, a_m)$ y $B = (b_{ij})$. Denotemos por $B_1, \dots, B_m$ las filas de B y por $B^1, \dots, B^n$ sus columnas. Entonces

$$\begin{aligned} RB &= (R \cdot B^1, R \cdot B^2, \dots, R \cdot B^n) = \\ &= (a_1 b_{11} + a_2 b_{21} + \cdots + a_m b_{m1}, a_1 b_{12} + a_2 b_{22} + \cdots + a_m b_{m2}, \dots, a_1 b_{1n} + a_2 b_{2n} + \cdots + a_m b_{mn}) = \\ &= a_1 (b_{11}, b_{12}, \dots, b_{1n}) + a_2 (b_{21}, b_{22}, \dots, b_{2n}) + \cdots + a_m (b_{m1}, b_{m2}, \dots, b_{mn}) = \\ &= a_1 B_1 + a_2 B_2 + \cdots + a_m B_m \end{aligned}$$

y RB es combinación lineal de las filas de B, como se pretendía.

b) Las filas de AB son $R_1 B$, donde R_1 es la fila i-ésima de A. De acuerdo con la parte *a*), cada fila de AB está en el espacio fila de B. De este modo, f-lin $AB \subseteq$ f-lin B, como se pretendía.

c) Usando la parte *b*),

$$\text{c-lin } AB = \text{f-lin } (AB)^T = \text{f-lin } B^T A^T \subseteq \text{f-lin } A^T = \text{c-lin } A$$

d) El espacio fila de AB está contenido en el de B; por tanto, rango $AB \leq$ rango B. Además, el espacio columna de AB está contenido en el de A; por consiguiente, rango $AB \leq$ rango A.

5.55. Sea A una matriz n-cuadrada. Probar que A es invertible si y sólo si rango $A = n$.

Nótese que las filas de la matriz identidad n-cuadrada I_n son linealmente independientes ya que I_n está en forma escalonada; por tanto, rango $I_n = n$. Ahora bien, si A es invertible, es equivalente por filas a I_n; de aquí rango $A = n$. Pero si A no es invertible, es equivalente por filas a una matriz con una fila nula, con lo que rango $A < n$. Esto es, A es invertible si y sólo si rango $A = n$.

APLICACIONES A LAS ECUACIONES LINEALES

5.56. Hallar la dimensión y una base del espacio solución W del sistema

$$\begin{aligned} x + 2y + z - 3t &= 0 \\ 2x + 4y + 4z - t &= 0 \\ 3x + 6y + 7z + t &= 0 \end{aligned}$$

Reducimos el sistema a forma escalonada:

$$\begin{aligned} x + 2y + z - 3t &= 0 \\ 2z + 5t &= 0 \\ 4z + 10t &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y + z - 3t &= 0 \\ 2z + 5t &= 0 \end{aligned}$$

Las variables libres son y y t y $\dim W = 2$. Tomamos:

i) $y = 1$, $z = 0$ para obtener la solución $u_1 = (-2, 1, 0, 0)$.

ii) $y = 0$, $t = 2$ para obtener la solución $u_2 = (11, 0, -5, 2)$.

Entonces $\{u_1, u_2\}$ es una base de W. [La elección $y = 0$, $t = 1$ en ii) introduciría fracciones en la solución.]

5.57. Encontrar un sistema homogéneo cuyo conjunto solución W esté generado por

$$\{(1, -2, 0, 3), (1, -1, -1, 4), (1, 0, -2, 5)\}$$

Sea $v = (x, y, x, t)$. Construimos la matriz M cuyas primeras filas son los vectores dados y cuya fila es v y la reducimos a forma escalonada:

$$M = \begin{pmatrix} 1 & -2 & 0 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 0 & -2 & 5 \\ x & y & z & t \end{pmatrix} \text{ a } \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & -2 & 2 \\ 0 & 2x + y & z & -3x + t \end{pmatrix} \text{ a } \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2x + y + z & -5x - y + t \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Las tres primeras filas originales muestran que W tiene dimensión 2, de modo que $v \in W$ si y sólo si la fila adicional no incrementa la dimensión del espacio fila. De acuerdo con ello, igualamos a 0 las dos últimas entradas de la tercera fila a la derecha para conseguir el sistema homogéneo requerido

$$\begin{aligned} 2x + y + z \quad &= 0 \\ 5x + y \quad - t &= 0 \end{aligned}$$

5.58. Sean $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ las variables libres de un sistema de ecuaciones lineales con n incógnitas homogéneo. Sea v_j la solución para la que $x_{i_j} = 1$, siendo todas las demás variables libres iguales a 0. Probar que las soluciones $v_1, v_2, \ldots, v_k$ son linealmente independientes.

Sea A la matriz cuyas filas son los v_i, respectivamente. Intercambiamos las columnas 1 e i_1, luego las 2 e $i_1, \ldots,$ y después las k e i_k, obteniendo la matriz $k \times n$

$$B = (I, C) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & c_{1,k+1} & \cdots & c_{1n} \\ 0 & 1 & 0 & \ldots & 0 & 0 & c_{2,k+1} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & c_{k,k+1} & \cdots & c_{kn} \end{pmatrix}$$

La matriz anterior B está en forma escalonada y por tanto sus filas son independientes; de aquí rango $B = k$. Como A y B son equivalentes por columnas, tienen el mismo rango, o sea, rango $A = k$. Pero A tiene k filas; por consiguiente, estas filas, es decir, los v_i, son linealmente independientes, como se pretendía.

5.59. Demostrar el Teorema 5.20.

Supongamos que $u_1, u_2, \ldots, u_r$ forman una base del espacio columna de A. (Existen r vectores tales, porque rango $A = r$.) Según el Teorema 5.19, cada sistema $AX = u_i$ tiene una solución, digamos v_i. De aquí

$$Av_1 = u_1,\ Av_2 = u_2,\ \ldots,\ Av_r = u_r \qquad [1]$$

Supongamos que dim $W = s$ y que $w_1, w_2, \ldots, w_s$ forman una base de W. Sea

$$B = \{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_s\}$$

Afirmamos que B es una base de K^n. Para poder hacerlo debemos probar que B genera K^n y es linealmente independiente.

a) *Demostración de que B genera K^n.* Supongamos que $v \in K^n$ y $Av = u$. En tal caso, $u = Av$ pertenece al espacio columna de A y necesariamente es una combinación lineal de los u_i. Por ejemplo,

$$Av = k_1u_1 + k_2u_2 + \cdots + k_ru_r \qquad [2]$$

Sea $v' = v - k_1v_1 - k_2v_2 - \cdots - k_rv_r$. Usando [1] y [2],

$$\begin{aligned} A(v') &= A(v - k_1v_1 - k_2v_2 - \cdots - k_rv_r) = \\ &= Av - k_1Av_1 - k_2Av_2 - \cdots - k_rAv_r = \\ &= Av - k_1u_1 - k_2u_2 - \cdots - k_ru_r = Av - Av = 0 \end{aligned}$$

Así v' pertenece a la solución W, por lo que es combinación lineal de los w_i. Por ejemplo, $v' = c_1w_1 + c_2w_2 + \cdots + c_sw_s$. Entonces

$$v = v' + \sum_{i=1}^{r} k_iv_i = \sum_{i=1}^{r} k_iv_i + \sum_{j=1}^{s} c_jw_j$$

De este modo, v es combinación lineal de los elementos de B y por ende B genera K^n.

b) *Demostración de que B es linealmente independiente.* Supongamos

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r + b_1w_1 + b_2w_2 + \cdots + b_sw_s = 0 \qquad [3]$$

Como $w_j \in W$, cada $Aw_j = 0$. Usando este hecho, junto a [1] y [3],

$$0 = A(0) = A\left(\sum_{i=1}^{r} a_iv_i + \sum_{j=1}^{s} b_jw_j\right) = \sum_{i=1}^{r} a_iAv_i + \sum_{j=1}^{s} b_jAw_j =$$

$$= \sum_{i=1}^{r} a_iu_i + \sum_{j=1}^{s} b_j0 = a_1u_1 + a_2u_2 + \cdots + a_ru_r$$

Dado que $u_1, \ldots, u_r$ son linealmente independientes, cada $a_i = 0$. Sustituyendo esto en [3] llegamos a

$$b_1w_1 + b_2w_2 + \cdots + b_sw_s = 0$$

No obstante, $w_1, \ldots, w_s$ son linealmente independientes, por lo que cada $b_j = 0$. Por esta razón, B es linealmente independiente.

De acuerdo con esto, B es una base de K^n. Dado que B tiene $r + s$ elementos, tenemos que $r + s = n$. En consecuencia, $\dim W = s = n - r$, como se pretendía.

SUMAS. SUMAS DIRECTAS. INTERSECCIONES

5.60. Sean U y W subespacios de un espacio vectorial V. Probar que: *a*) U y W están contenidos en $U + W$; *b*) $U + W$ es el menor subespacio de V que contiene U y W, esto es, $U + W = \text{lin}(U, W)$, envolvente lineal de U y W; *c*) $W + W = W$.

a) Sea $u \in U$. Por hipótesis, W es un subespacio de V, por lo que $0 \in W$. Por consiguiente, $u = u + 0 \in U + W$. Según esto, U está contenido en $U + W$. De forma similar, W está contenido en $U + W$.

b) Dado que $U + W$ es un subespacio de V que contiene tanto U como W, debe contener también la envolvente lineal de U y W, o sea, $\text{lin}(U, W) \subseteq U + W$.

Por otra parte, si $v \in U + W$, necesariamente $v = u + w = 1u + 1w$, donde $u \in U$ y $w \in W$; de aquí que v sea combinación lineal de elementos en $U \cup W$ y por tanto pertenezca a $\text{lin}(U, W)$. En consecuencia, $U + W \subseteq \text{lin}(U, W)$.

c) Como W es un subespacio de V, tenemos que W es cerrado bajo la suma vectorial; por este motivo, $W + W \subseteq W$. Por la parte *a*), $W \subseteq W + W$. Entonces $W + W = W$.

5.61. Dar un ejemplo de un subconjunto S de $\mathbf{R}^2$ tal que: *a*) $S + S \subset S$ (contenido propiamente); *b*) $S \subset S + S$ (contenido propiamente); *c*) $S + S = S$ pero S no es un subespacio de $\mathbf{R}^2$

a) Sea $S = \{(0, 5), (0, 6), (0, 7), \ldots\}$. En tal caso $S + S \subset S$.

b) Sea $S = \{(0, 0), (0, 1)\}$. En tal caso $S \subset S + S$.

c) Sea $S = \{(0, 0), (0, 1), (0, 2), (0, 3), \ldots\}$. En tal caso $S + S = S$.

5.62. Supóngase que U y W son subespacios 4-dimensionales distintos de un espacio vectorial V con $\dim V = 6$. Hallar todas las dimensiones posibles de $U \cap W$.

Como U y W son distintos, $U + W$ contiene propiamente U y W; consecuentemente, $\dim(U + W) > 4$. Pero $\dim(U + W)$ no puede ser mayor que 6, ya que $\dim V = 6$. Por tanto, tenemos dos posibilidades: i) $\dim(U + W) = 5$, o ii) $\dim(U + W) = 6$. Por el Teorema 5.22,

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) = 8 - \dim(U + W)$$

Así, i) $\dim(U \cap W) = 3$, o ii) $\dim(U \cap W) = 2$.

5.63. Considérense los siguientes subespacios de $\mathbf{R}^4$:

$$U = \text{lin}\{(1, 1, 0, -1), (1, 2, 3, 0), (2, 3, 3, -1)\}$$
$$W = \text{lin}\{(1, 2, 2, -2), (2, 3, 2, -3), (1, 3, 4, -3)\}$$

Hallar: *a*) $\dim(U + W)$, *b*) $\dim(U \cap W)$.

a) $U + W$ es el espacio generado por los seis vectores. Construimos, pues, la matriz cuyas filas son los seis vectores dados y la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \\ 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \sim \begin{pmatrix} & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 2 & 4 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Como la matriz escalonada tiene tres filas no nulas, $\dim(U + W) = 3$.

b) Hallamos primero $\dim U$ y $\dim W$. Formamos las dos matrices cuyas filas son los generadores de U y W, respectivamente, y las reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

y

$$\begin{pmatrix} 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 1 & 2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dado que cada una de las matrices escalonadas tiene dos filas no nulas, $\dim U = 2$ y $\dim W = 2$. Usando el Teorema 5.22, esto es, $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, tenemos

$$3 = 2 + 2 - \dim(U \cap W) \qquad \text{o} \qquad \dim(U \cap W) = 1$$

5.64. Sean U y W los siguientes subespacios de $\mathbf{R}^4$:

$$U = \{(a, b, c, d): b + c + d = 0\}, \qquad W = \{(a, b, c, d): a + b = 0, c = 2d\}$$

Encontrar una base y la dimensión de: *a*) U, *b*) W, *c*) $U \cap W$, *d*) $U + W$.

a) Buscamos una base del conjunto de soluciones (a, b, c, d) de la ecuación

$$b + c + d = 0 \qquad \text{o} \qquad 0 \cdot a + b + c + d = 0$$

Las variables libres son a, c y d. Tomemos:

1. $a = 1$, $c = 0$, $d = 0$ para obtener la solución $v_1 = (1, 0, 0, 0)$.
2. $a = 0$, $c = 1$, $d = 0$ para obtener la solución $v_2 = (0, -1, 1, 0)$.
3. $a = 0$, $c = 0$, $d = 1$ para obtener la solución $v_3 = (0, -1, 0, 1)$.

El conjunto $\{v_1, v_2, v_3\}$ es una base de U y $\dim U = 3$.

b) Buscamos una base del conjunto de soluciones (a, b, c, d) del sistema

$$\begin{aligned} a + b &= 0 \\ c &= 2d \end{aligned} \qquad \text{o} \qquad \begin{aligned} a + \ b &= 0 \\ c - 2d &= 0 \end{aligned}$$

Las variables libres son b y d. Tomemos:

1. $b = 1$, $d = 0$ para obtener la solución $v_1 = (-1, 1, 0, 0)$
2. $b = 0$, $d = 1$ para obtener la solución $v_2 = (0, 0, 2, 1)$

El conjunto $\{v_1, v_2\}$ es una base de W y $\dim W = 2$.

c) $U \cap W$ consiste en aquellos vectores (a, b, c, d) que satisfacen las condiciones que definen U y las que definen W, es decir, las tres ecuaciones

$$\begin{aligned} b + c + d &= 0 \\ a + b \qquad &= 0 \\ c \qquad &= 2d \end{aligned} \qquad \text{o} \qquad \begin{aligned} a + b \qquad\quad &= 0 \\ b + c + \ d &= 0 \\ c - 2d &= 0 \end{aligned}$$

La variable libre es d. Tomemos $d = 1$ para obtener la solución $v = (3, -3, 2, 1)$. De este modo, $\{v\}$ es una base de $U \cap W$ y $\dim(U \cap W) = 1$.

d) Según el Teorema 5.22,

$$\dim (U + W) = \dim U + \dim W - \dim (U \cap W) = 3 + 2 - 1 = 4$$

Así $U + W = \mathbf{R}^4$. De acuerdo con esto, cualquier base de $\mathbf{R}^4$, por ejemplo la base usual, será una base de $U + W$.

5.65. Considérense los siguientes subespacios de $\mathbf{R}^5$:

$$U = \text{lin}\,\{(1, 3, -2, 2, 3), (1, 4, -3, 4, 2), (2, 3, -1, -2, 9)\}$$
$$W = \text{lin}\,\{(1, 3, 0, 2, 1), (1, 5, -6, 6, 3), (2, 5, 3, 2, 1)\}$$

Hallar una base y la dimensión de *a*) $U + W$, *b*) $U \cap W$.

a) $U + W$ es el espacio generado por los seis vectores. Por tanto, formamos la matriz cuyas filas son los seis vectores dados y después la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & -4 & 4 & 0 \\ 0 & -1 & 7 & -2 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 6 & 0 & -6 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

El conjunto de filas no nulas de la matriz escalonada,

$$\{(1, 3, -2, 2, 3), (0, 1, -1, 2, -1), (0, 0, 2, 0, -2)\}$$

es una base de $U + W$, de modo que $\dim (U + W) = 3$.

b) Hallamos primero los sistemas homogéneos cuyos conjuntos solución sean U y W, respectivamente. Construimos la matriz cuyas tres primeras filas generen U y cuya última fila sea (x, y, z, s, t). Luego la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ x & y & z & s & t \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & -3x+y & 2x+z & -2x+s & -3x+t \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & -x+y+z & 4x-2y+s & -6x+y+t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Igualamos a cero las entradas de la tercera fila para conseguir el sistema homogéneo cuyo espacio solución es U:

$$-x+y+z=0 \qquad 4x-2y+s=0 \qquad -6x+y+t=0$$

Ahora formamos la matriz cuyas primeras filas generen W, siendo la última (x, y, z, s, t). Acto seguido, la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \\ x & y & z & s & t \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 2 & -6 & 4 & 2 \\ 0 & -1 & 3 & -2 & -1 \\ 0 & -3x+y & z & -2x+s & -x+t \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 1 & -3 & 2 & 1 \\ 0 & 0 & -9x+3y+z & 4x-2y+s & 2x-y+t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Igualamos a cero las entradas de la tercera fila para conseguir el sistema homogéneo cuyo espacio solución es W:

$$-9x+3y+z=0 \qquad 4x-2y+s=0 \qquad 2x-y+t=0$$

Combinamos los dos sistemas precedentes para obtener un sistema homogéneo cuyo espacio solución es $U \cap W$ y lo resolvemos:

$$\begin{cases} -x+y+z=0 \\ 4x-2y+s=0 \\ -6x+y+t=0 \\ -9x+3y+z=0 \\ 4x-2y+s=0 \\ 2x-y+t=0 \end{cases} \quad \text{o} \quad \begin{cases} -x+y+z=0 \\ 2y+4z+s=0 \\ -5y-6z+t=0 \\ -6y-8z=0 \\ 2y+4z+s=0 \\ y+2z+t=0 \end{cases}$$

$$\text{o} \quad \begin{cases} -x+y+z=0 \\ 2y+4z+s=0 \\ 8z+5s+2t=0 \\ 4z+3s=0 \\ s-2t=0 \end{cases} \quad \text{o} \quad \begin{cases} -x+y+z=0 \\ 2y+4z+s=0 \\ 8z+5s+2t=0 \\ s-2t=0 \end{cases}$$

Hay una variable libre, que es t; por consiguiente, $\dim(U \cap W) = 1$. Haciendo $t = 2$ llegamos a la solución $x = 1$, $y = 4$, $z = -3$, $s = 4$, $t = 2$. Siendo así, $\{(1, 4, -3, 4, 2)\}$ es una base de $U \cap W$.

5.66. Sean U y W los subespacios de $\mathbf{R}^3$ definidos según

$$U = \{(a, b, c): a = b = c\} \qquad \text{y} \qquad W = \{(0, b, c)\}$$

(Nótese que W es el plano yz.) Probar que $\mathbf{R}^3 = U \oplus W$.

Adviértase primero que $U \cap W = \{0\}$, para $v = (a, b, c) \in U \cap W$ implica que

$$a = b = c \quad \text{y} \quad a = 0 \quad \text{lo que implica que} \quad a = 0,\ b = 0,\ c = 0$$

Afirmamos también que $\mathbf{R}^3 = U + W$. Ello es debido a que si $v = (a, b, c) \in \mathbf{R}^3$,

$$v = (a, a, a) + (0, b - a, c - a) \quad \text{donde} \quad (a, a, a) \in U \quad \text{y} \quad (0, b - a, c - a) \in W$$

Ambas condiciones, $U \cap W = \{0\}$ y $\mathbf{R}^3 = U + W$, implican $\mathbf{R}^3 = U \oplus W$.

5.67. Sea V el espacio vectorial de las matrices n-cuadradas sobre un cuerpo K.

a) Probar que $V = U \oplus W$, donde U y W son los subespacios de las matrices simétricas y antisimétricas, respectivamente. (Recuérdese que M es simétrica si y sólo si $M = M^T$, y antisimétrica si y sólo si $M^T = -M$.)

b) Probar que $V \neq U \oplus W$, donde U y W son los subespacios de las matrices triangulares superiores e inferiores, respectivamente. (Nótese que $V = U + W$.)

a) Comencemos por demostrar que $V = U + W$. Sea A una matriz n-cuadrada arbitraria. Adviértase que

$$A = \tfrac{1}{2}(A + A^T) + \tfrac{1}{2}(A - A^T)$$

Afirmamos que $\frac{1}{2}(A + A^T) \in U$ y que $\frac{1}{2}(A - A^T) \in W$. Esto se debe a que

$$(\tfrac{1}{2}(A + A^T))^T = \tfrac{1}{2}(A + A^T)^T = \tfrac{1}{2}(A^T + A^{TT}) = \tfrac{1}{2}(A + A^T)$$

o sea, $\frac{1}{2}(A + A^T)$ es simétrica. Además,

$$(\tfrac{1}{2}(A - A^T))^T = \tfrac{1}{2}(A - A^T)^T = \tfrac{1}{2}(A^T - A) = -\tfrac{1}{2}(A - A^T)$$

esto es, $\frac{1}{2}(A - A^T)$ es antisimétrica.

A continuación mostremos que $U \cap W = \{0\}$. Supongamos que $M \in U \cap W$. En ese caso, $M = M^T$ y $M^T = -M$, lo que implica $M = -M$ y por tanto $M = 0$. Así $U \cap W = \{0\}$. En consecuencia, $V = U \oplus W$.

b) $U \cap W \neq \{0\}$ porque $U \cap W$ está constituido por todas las matrices diagonales. La suma no puede, pues, ser directa.

5.68. Supóngase que U y W son subespacios de un espacio vectorial V, que $S = \{u_i\}$ genera U y que $T = \{w_j\}$ genera W. Demostrar que $S \cup T$ genera $U + W$. (De acuerdo con ello, por inducción, si S_i genera W_i para $i = 1, 2, \ldots, n$, $S_1 \cup \cdots \cup S_n$ genera $W_1 + \cdots + W_n$.)

Sea $v \in U + W$. Entonces $v = u + w$, donde $u \in U$ y $w \in W$. Puesto que S genera U, u es combinación lineal de los u_i, y dado que T genera W, w es combinación lineal de los w_j:

$$u = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n} \qquad a_j \in K$$
$$w = b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m} \qquad b_j \in K$$

De este modo,

$$v = u + w = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n} + b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m}$$

En consecuencia, $S \cup T = \{u_i, v_j\}$ genera $U + W$.

5.69. Demostrar el Teorema 5.22.

Observemos que $U \cap W$ es subespacio tanto de U como de W. Supongamos que $\dim U = m$, $\dim W = n$ y $\dim U \cap W = r$. Supongamos, asimismo, que $\{v_1, \ldots, v_r\}$ es una base de $U \cap W$. Según el Teorema 5.6, podemos extender $\{v_i\}$ a una base de U y a una de W; por ejemplo,

$$\{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}\} \quad \text{y} \quad \{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$$

son bases de U y W, respectivamente. Sea

$$B = \{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}, w_1, \ldots, w_{n-r}\}$$

Nótese que B tiene exactamente $m + n - r$ elementos. El teorema quedará demostrado si somos capaces de probar que B es una base de $U + W$. Como $\{v_i, u_j\}$ genera U y $\{v_i, w_k\}$ genera W, la unión $B = \{v_i, u_j, w_k\}$ generará $U + W$. Por tanto, basta demostrar que B es independiente.

Supongamos que

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0 \qquad [1]$$

donde a_i, b_j, c_k son escalares. Sea

$$v = a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} \qquad [2]$$

Por [1] tenemos también que

$$v = -c_1w_1 - \cdots - c_{n-r}w_{n-r} \qquad [3]$$

Como $\{v_i, u_j\} \subseteq U$, $v \in U$ por [2]; y como $\{w_k\} \subseteq W$, $v \in W$ por [3]. En consecuencia, $v \in U \cap W$. Ahora bien, $\{v_i\}$ es una base de $U \cap W$, por lo que existen escalares $d_1, \ldots, d_r$, para los que $v = d_1v_1 + \cdots + d_rv_r$. Así por [3] tenemos

$$d_1v_1 + \cdots + d_rv_r + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0$$

Pero $\{v_i, w_k\}$ es una base de W y necesariamente es independiente. De aquí que la ecuación anterior haga forzoso $c_1 = 0, \ldots, c_{n-r} = 0$. Sustituyéndolo en [1],

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} = 0$$

Pero $\{v_i, u_j\}$ es una base de U y por tanto es independiente. De aquí que la ecuación anterior haga forzoso $a_1 = 0, \ldots, a_r = 0$, $b_1 = 0, \ldots, b_{m-r} = 0$.

Dado que la ecuación [1] implica que los a_i, b_j y c_k son todos 0, $B = \{v_i, u_j, w_k\}$ es independiente y queda probado el teorema.

5.70. Demostrar el Teorema 5.23.

Supongamos $V = U \oplus W$. Cualquier vector $v \in V$ puede, pues, escribirse de forma única como $v = u + w$, con $u \in U$ y $w \in W$. Así, en particular, $V = U + W$. Supongamos ahora que $v \in U \cap W$. Entonces:

1. $v = v + 0$, donde $v \in U$, $0 \in W$.
2. $v = 0 + v$; donde $0 \in U$, $v \in W$.

Como tal suma para v debe ser única, $v = 0$. De acuerdo con ello, $U \cap W = \{0\}$.

Por otra parte, supongamos que $V = U + W$ y $U \cap W = \{0\}$. Sea $v \in V$. Dado que $V = U + W$, existen $u \in U$ y $w \in W$ tales que $v = u + w$. Debemos probar que la suma es única. Para ello, supongamos que $v = u' + w'$ también, donde $u' \in U$ y $w' \in W$. En tal caso,

$$u + w = u' + w' \quad \text{y por tanto} \quad u - u' = w' - w$$

Pero $u - u' \in U$ y $w' - w \in W$; por consiguiente, como $U \cap W = \{0\}$,

$$u - u' = 0, \ w' - w = 0 \qquad \text{y} \qquad u = u', \ w = w'$$

De este modo, tal suma para $v \in V$ es única y $V = U \oplus W$.

5.71. Demostrar el Teorema 5.24 (para dos sumandos): Supongamos que $V = U \oplus W$. Supongamos que $S = \{u_1, \ldots, u_m\}$ y $S' = \{w_1, \ldots, w_n\}$ son subconjuntos linealmente independientes de U y W, respectivamente. En tal caso: *a*) $S \cup S'$ es linealmente independiente en V; *b*) si S es una base de U y S' una de W, $S \cup S'$ es base de V; *c*) $\dim V = \dim U + \dim W$.

a) Supongamos $a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n = 0$, donde a_i, b_j son escalares. Entonces

$$0 = (a_1 u_1 + \cdots + a_m u_m) + (b_1 w_1 + \cdots + b_n w_n) = 0 + 0$$

donde $0, a_1 u_1 + \cdots + a_m u_m \in U$ y $0, b_1 w_1 + \cdots + b_n w_n \in W$. Puesto que tal suma para 0 es única, esto nos conduce a

$$a_1 u_1 + \cdots + a_m u_m = 0 \qquad b_1 w_1 + \cdots + b_n w_n = 0$$

Como S es linealmente independiente, cada $a_i = 0$, y como S' es linealmente independiente, cada $b_j = 0$. Siendo así, $S \cup S'$ es linealmente independiente.

b) Por la parte *a*) sabemos que $S \cup S'$ es linealmente independiente, y por el Problema 5.68 que $S \cup S'$ genera V. Por tanto, $S \cup S'$ es una base de V.

c) Se obtiene directamente de la parte *b*).

VECTORES COORDENADOS

5.72. Sea S la base de $\mathbf{R}^2$ consistente en $u_1 = (2, 1)$ y $u_2 = (1, -1)$. Hallar el vector coordenado $[v]$ de v relativo a S, siendo $v = (a, b)$.

Tomamos $v = xu_1 + yu_2$ para obtener $(a, b) = (2x + y, x - y)$. Resolvemos $2x + y = a$ y $x - y = b$ para obtener $x = (a + b)/3$, $y = (a - 2b)/3$. Así $[v] = [(a + b)/3, (a - 2b)/3]$.

5.73. Considérese el espacio vectorial $\mathbf{P}_3(t)$ de los polinomios reales en t de grado ≤ 3.

a) Probar que $S = \{1, 1 - t, (1 - t)^2, (1 - t)^3\}$ es una base de $\mathbf{P}_3(t)$.

b) Hallar el vector coordenado $[u]$ de $u = 2 - 3t + 3t^2 + 2t^3$ respecto a S.

a) El grado de $(1 - t)^k$ es k, luego ningún polinomio en S es combinación lineal de los precedentes. Los polinomios son, pues, linealmente independientes y, dado que $\dim \mathbf{P}_3(t) = 4$, forman una base de $\mathbf{P}_3(t)$.

b) Tomemos u como combinación lineal de los vectores de la base utilizando incógnitas x, y, z, s:

$$\begin{aligned} u = 2 - 3t + t^2 + 2t^3 &= x(1) + y(1 - t) + z(1 - t)^2 + s(1 - t)^3 = \\ &= x(1) + y(1 - t) + z(1 - 2t + t^2) + s(1 - 3t + 3t^2 - t^3) = \\ &= x + y - yt + z - 2zt + zt^2 + s - 3st + 3st^2 - st^3 = \\ &= (x + y + z + s) + (-y - 2z - 3s)t + (z + 3s)t^2 + (-s)t^3 \end{aligned}$$

A continuación igualamos entre sí los coeficientes de las mismas potencias de t:

$$x + y + z + s = 2 \qquad -y - 2z - 3s = -3 \qquad z + 3s = 1 \qquad -s = 2$$

Resolviendo, $x = 2$, $y = -5$, $z = 7$, $s = -2$. Así $[u] = [2, -5, 7, -2]$.

5.74. Considérese la matriz $A = \begin{pmatrix} 2 & 3 \\ 4 & -7 \end{pmatrix}$ en el espacio vectorial V de las matrices reales 2×2. Determinar el vector coordenado $[A]$ de la matriz A respecto a

$$\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

la base usual de V.

Tenemos

$$\begin{pmatrix} 2 & 3 \\ 4 & -7 \end{pmatrix} = x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + z\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + t\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

De este modo, $x = 2$, $y = 3$, $z = 4$, $t = -7$. Por tanto, $[A] = [2, 3, 4, -7]$, cuyas componentes son los elementos de A escritos fila a fila.

Nota: El resultado anterior es válido en general, es decir, si A es cualquier matriz en el espacio vectorial de las matrices $m \times n$ sobre un cuerpo K, el vector coordenado $[A]$ de A relativo a la base usual de V es el vector coordenado mn de K^{mn} cuyas componentes son los elementos de A escritos fila por fila.

CAMBIO DE BASE

En esta sección se representará un vector $v \in V$ respecto a una base S de V por su vector columna coordenado,

$$[v]_S = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = [a_1, a_2, \dots, a_n]^T$$

(que es una matriz $n \times 1$).

5.75. Considérense las siguientes bases de $\mathbf{R}^2$:

$$S_1 = \{u_1 = (1, -2), u_2 = (3, -4)\} \quad \text{y} \quad S_2 = \{v_1 = (1, 3), v_2 = (3, 8)\}$$

a) Encontrar las coordenadas de un vector arbitrario $v = (a, b)$ relativas a la base $S_1 = \{u_1, u_2\}$.

b) Hallar la matriz de cambio de base P desde S_1 hasta S_2.

c) Encontrar las coordenadas de un vector arbitrario $v = (a, b)$ relativas a la base $S_2 = \{v_1, v_2\}$.

d) Determinar la matriz de cambio de base Q desde S_2 hasta S_1.

e) Comprobar que $Q = P^{-1}$.

f) Mostrar que $P[v]_{S_2} = [v]_{S_1}$ para todo vector $v = (a, b)$. (Véase el Teorema 5.27.)

g) Mostrar que $P^{-1}[v]_{S_1} = [v]_{S_2}$ para todo vector $v = (a, b)$. (Véase el Teorema 5.27.)

a) Sea $v = xu_1 + yu_2$ para incógnitas x e y:

$$\begin{pmatrix} a \\ b \end{pmatrix} = x\begin{pmatrix} 1 \\ -2 \end{pmatrix} + y\begin{pmatrix} 3 \\ -4 \end{pmatrix} \quad \text{o} \quad \begin{aligned} x + 3y &= a \\ -2x - 4y &= b \end{aligned} \quad \text{o} \quad \begin{aligned} x + 3y &= a \\ 2y &= 2a + b \end{aligned}$$

Despejamos x e y en términos de a y b para llegar a $x = -2a - \frac{3}{2}b$, $y = a + \frac{1}{2}b$. Así pues,

$$(a, b) = (-2a - \tfrac{3}{2}b)u_1 + (a + \tfrac{1}{2}b)u_2 \quad \text{o} \quad [(a, b)]_{S_1} = [-2a - \tfrac{3}{2}b, a + \tfrac{1}{2}b]^T$$

b) Usamos *a*) para escribir cada uno de los vectores v_1 y v_2 de la base S_2 como combinación lineal de los vectores u_1 y u_2 de S_1:

$$v_1 = (1, 3) = (-2 - \tfrac{9}{2})u_1 + (1 + \tfrac{3}{2})u_2 = (-\tfrac{13}{2})u_1 + (\tfrac{5}{2})u_2$$

$$v_2 = (3, 8) = (-6 - 12)u_1 + (3 + 4)u_2 = -18u_1 + 7u_2$$

Entonces P es la matriz cuyas columnas son las coordenadas de v_1 y v_2 respecto a la base S_1, o sea,

$$P = \begin{pmatrix} -\frac{13}{2} & -18 \\ \frac{5}{2} & 7 \end{pmatrix}$$

c) Sea $v = xv_1 + yv_2$ para escalares desconocidos x e y:

$$\begin{pmatrix} a \\ b \end{pmatrix} = x\begin{pmatrix} 1 \\ 3 \end{pmatrix} + y\begin{pmatrix} 3 \\ 8 \end{pmatrix} \quad \text{o} \quad \begin{aligned} x + 3y &= a \\ 3x + 8y &= b \end{aligned} \quad \text{o} \quad \begin{aligned} x + 3y &= a \\ -y &= b - 3a \end{aligned}$$

Resolvemos para x e y llegando a $x = -8a + 3b$, $y = 3a - b$. Así

$$(a, b) = (-8a + 3b)v_1 + (3a - b)v_2 \quad \text{o} \quad [(a, b)]S_2 = [-8a + 3b, 3a - b]^T$$

d) Usamos *c*) para escribir cada uno de los vectores u_1 y u_2 de la base S_1 como combinación lineal de los vectores v_1 y v_2 de S_2:

$$u_1 = (1, -2) = (-8 - 6)v_1 + (3 + 2)v_2 = -14v_1 + 5v_2$$

$$u_2 = (3, -4) = (-24 - 12)v_1 + (9 + 4)v_2 = -36v_1 + 13v_2$$

Escribimos las coordenadas de u_1 y u_2 respecto a S_2 como columnas obteniendo $Q = \begin{pmatrix} -14 & -36 \\ 5 & 13 \end{pmatrix}$.

e)

$$QP = \begin{pmatrix} -14 & -36 \\ 5 & 13 \end{pmatrix}\begin{pmatrix} -\frac{13}{2} & -18 \\ \frac{5}{2} & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

f) Utilizamos *a*), *b*) y *c*) para obtener

$$P[v]_{S_2} = \begin{pmatrix} -\frac{13}{2} & -18 \\ \frac{5}{2} & 7 \end{pmatrix}\begin{pmatrix} -8a + 3b \\ 3a - b \end{pmatrix} = \begin{pmatrix} -2a - \frac{3}{2}b \\ a + \frac{1}{2}b \end{pmatrix} = [v]_{S_1}$$

g) Utilizamos *a*), *c*) y *d*) para obtener

$$P^{-1}[v]_{S_1} = Q[v]_{S_1} = \begin{pmatrix} -14 & -36 \\ 5 & 13 \end{pmatrix}\begin{pmatrix} -2a - \frac{3}{2}b \\ a + \frac{1}{2}b \end{pmatrix} = \begin{pmatrix} -8a + 3b \\ 3a - b \end{pmatrix} = [v]_{S_2}$$

5.76. Supóngase que los siguientes vectores constituyen una base S de K^n:

$$v_1 = (a_1, a_2, \ldots, a_n), \qquad v_2 = (b_1, b_2, \ldots, b_n), \qquad \ldots, \qquad v_n = (c_1, c_2, \ldots, c_n)$$

Probar que la matriz de cambio de base desde la base usual $E = \{e_i\}$ de K^n hasta la base S es la matriz P cuyas columnas son los vectores $v_1, v_2, \ldots, v_n$, respectivamente.

Como $e_1, e_2, \ldots, e_n$ forman la base usual E de K^n,

$$\begin{aligned} v_1 &= (a_1, a_2, \ldots, a_n) = a_1e_1 + a_2e_2 + \cdots + a_ne_n \\ v_2 &= (b_1, b_2, \ldots, b_n) = b_1e_1 + b_2e_2 + \cdots + b_ne_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_n &= (c_1, c_2, \ldots, c_n) = c_1e_1 + c_2e_2 + \cdots + c_ne_n \end{aligned}$$

Escribiendo las coordenadas como columnas tendremos

$$P = \begin{pmatrix} a_1 & b_1 & \ldots & c_1 \\ a_2 & b_2 & \ldots & c_2 \\ \ldots & \ldots & \ldots & \ldots \\ a_n & b_n & \ldots & c_n \end{pmatrix}$$

como se pretendía.

5.77. Considérese la base $S = \{u_1 = (1, 2, 0), u_2 = (1, 3, 2), u_3 = (0, 1, 3)\}$ de $\mathbf{R}^3$. Hallar:

a) La matriz de cambio de base P desde la base usual $E = \{e_1, e_2, e_3\}$ de $\mathbf{R}^3$ hasta la base S.

b) La matriz de cambio de base Q desde la base anterior S regresando hasta la base usual E de $\mathbf{R}^3$.

a) Dado que E es la base usual, escribimos sencillamente los vectores de la base S como columnas:

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 2 & 3 \end{pmatrix}$$

b) **Método 1.** Expresamos cada vector de la base E como combinación lineal de los vectores de la base S, calculando primero las coordenadas de un vector arbitrario $v = (a, b, c)$ relativas a la base S. Tenemos:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + z\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \qquad \text{o} \qquad \begin{aligned} x + y \quad &= a \\ 2x + 3y + z &= b \\ 2y + 3z &= c \end{aligned}$$

Despejamos x, y, z llegando a $x = 7a - 3b + c$, $y = -6a + 3b - c$, $z = 4a - 2b + c$. Entonces

$$v = (a, b, c) = (7a - 3b + c)u_1 + (-6a + 3b - c)u_2 + (4a - 2b + c)u_3$$

o $$[v]_S = [(a, b, c)]_S = [7a - 3b + c, \ -6a + 3b - c, \ 4a - 2b + c]^T$$

Empleando la fórmula para $[v]_S$ precedente y escribiendo después las coordenadas de los e_i como columnas,

$$\begin{aligned} e_1 &= (1, 0, 0) = 7u_1 - 6u_2 + 4u_3 \\ e_2 &= (0, 1, 0) = -3u_1 + 3u_2 - 2u_3 \\ e_3 &= (0, 0, 1) = u_1 - u_2 + u_3 \end{aligned} \qquad \text{y} \qquad Q = \begin{pmatrix} 7 & -3 & 1 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{pmatrix}$$

Método 2. Calculamos P^{-1} reduciendo por filas $M = (P \mid I)$ a la forma $(I \mid P^{-1})$:

$$M = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -6 & 3 & -1 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 7 & -3 & 1 \\ 0 & 1 & 0 & -6 & 3 & 1 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right)$$

De este modo, $Q = P^{-1} = \begin{pmatrix} 7 & -3 & 1 \\ -6 & 3 & -1 \\ 4 & -2 & 1 \end{pmatrix}$.

5.78. Supóngase que los ejes x e y en el plano $\mathbf{R}^2$ se giran 45° en el sentido contrario al de las agujas de un reloj, de forma que el nuevo eje x' esté a lo largo de la recta $x = y$ y el nuevo eje y' a lo largo de la recta $x = -y$. Encontrar: *a*) la matriz de cambio de base P; *b*) las nuevas coordenadas del punto $A(5, 6)$ bajo la rotación dada.

a) Los vectores unitarios en las direcciones de los nuevos ejes x' e y' son, respectivamente,

$$u_1 = (\sqrt{2}/2, \sqrt{2}/2) \qquad \text{y} \qquad u_2 = (-\sqrt{2}/2, \sqrt{2}/2)$$

(Los vectores unitarios en las direcciones de los ejes originales x e y son los vectores de la base usual de $\mathbf{R}^2$.) Escribimos, pues, las coordenadas de u_1 y u_2 como columnas para obtener

$$P = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

b) Multiplicamos las coordenadas del punto por P^{-1}:

$$\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}\begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 11\sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}$$

[Dado que P es ortogonal, P^{-1} es simplemente la traspuesta de P.]

5.79. Considérense las bases $S = \{1, i\}$ y $S' = \{1 + i, 1 + 2i\}$ del cuerpo complejo $\mathbf{C}$ sobre el cuerpo real $\mathbf{R}$. Hallar: *a*) la matriz de cambio de base P desde la base S hasta la S'; *b*) la matriz del cambio de base Q desde la base S' volviendo hasta la S.

a) Tenemos

$$\begin{aligned} 1 + i &= 1(1) + 1(i) \\ 1 + 2i &= 1(1) + 2(i) \end{aligned} \qquad \text{y por tanto} \qquad P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

b) Utilizamos la fórmula para la inversa de una matriz 2×2 para conseguir $Q = P^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.

5.80. Supóngase que P es la matriz de cambio de base desde una base $\{u_i\}$ hasta otra $\{w_i\}$ y que Q es la matriz de cambio de base desde la $\{w_i\}$ hasta la $\{u_i\}$. Probar que P es invertible y que $Q = P^{-1}$.

Supongamos, para $i = 1, 2, \ldots, n$,

$$w_i = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j \qquad [1]$$

y, para $j = 1, 2, \ldots, n$,

$$u_j = b_{j1}w_1 + b_{j2}w_2 + \cdots + b_{jn}w_n = \sum_{k=1}^{n} b_{jk}w_k \qquad [2]$$

Sean $A = (a_{ij})$ y $B = (b_{jk})$. En ese caso, $P = A^T$ y $Q = B^T$. Sustituyendo [2] en [1]

$$w_1 = \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk}w_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)w_k$$

Puesto que $\{w_i\}$ es una base, $\sum a_{ij}b_{jk} = \delta_{ik}$, donde δ_{ik} es la delta de Kronecker, esto es, $\delta_{ik} = 1$ si $i = k$ pero $\delta_{ik} = 0$ si $i \neq k$. Supongamos que $AB = (c_{ik})$. Entonces $c_{ik} = \delta_{ik}$. De acuerdo con ello, $AB = I$, luego

$$QP = B^TA^T = (AB)^T = I^T = I$$

Así pues, $Q = P^{-1}$.

5.81. Demostrar el Teorema 5.27.

Supongamos que $S = \{u_1, \ldots, u_n\}$ y $S' = \{w_1, \ldots, w_n\}$, y que, para $i = 1, \ldots, n$,

$$w_i = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j$$

Por consiguiente, P es la matriz n-cuadrada cuya fila j-ésima es

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \qquad [1]$$

Asimismo, supongamos $v = k_1w_1 + k_2w_2 + \cdots + k_nw_n = \sum_{i=1}^{n} k_iw_i$. Entonces

$$[v]_{S'} = [k_{1j}, k_{2j}, \ldots, k_{nj}]^T \qquad [2]$$

Sustituyendo los w_i en la ecuación para v obtenemos

$$v = \sum_{i=1}^{n} k_iw_i = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}u_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)u_j =$$

$$= \sum_{j=1}^{n} (a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)u_j$$

En consecuencia, $[v]_S$ es el vector columna cuya entrada j-ésima es

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \qquad [3]$$

Por otra parte, la entrada j-ésima de $P[v]_{S'}$ se obtiene multiplicando la j-ésima fila de P por $[v]_{S'}$, o sea, [1] por [2]. Pero el producto de [1] y [2] es [3]; de aquí que $P[v]_{S'}$ y $[v]_S$ tengan las mismas entradas. Así $P[v]_{S'} = [v]_S$, como se pretendía.

Aún más, multiplicar lo anterior por P^{-1} conduce a $P^{-1}[v]_S = P^{-1}P[v]_{S'} = [v]_{S'}$.

PROBLEMAS VARIOS

5.82. Considérese una sucesión finita de vectores $S = \{v_1, v_2, \ldots, v_n\}$. Sea T la sucesión de vectores que se obtiene de S mediante una de las siguientes «operaciones elementales»: i) intercambiar dos vectores, ii) multiplicar un vector por un escalar no nulo, iii) sumar a un vector un múltiplo de otro. Probar que S y T generan el mismo espacio W. Probar también que T es independiente si y sólo si lo es S.

Observemos que, para cada operación, los vectores de T son combinaciones lineales de los de S. Por otra parte, cada operación tiene una inversa del mismo tipo. (Demuéstrese.) En consecuencia, los vectores de S son combinaciones lineales de los de T. De este modo, S y T generan el mismo espacio W. Además, T es independiente si y sólo si $\dim W = n$, lo que es cierto si y sólo si S también es independiente.

5.83. Sean $A = (a_{ij})$ y $B = (b_{ij})$ matrices $m \times n$ sobre un cuerpo K equivalentes por filas y $v_1, \ldots, v_n$ vectores cualesquiera en un espacio vectorial V sobre K. Sean

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \qquad & w_1 &= b_{11}v_1 + b_{12}v_2 + \cdots + b_{1n}v_n \\ u_2 &= a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & w_2 &= b_{21}v_1 + b_{22}v_2 + \cdots + b_{2n}v_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots & &\cdots\cdots\cdots\cdots\cdots\cdots \\ u_m &= a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n & w_m &= b_{m1}v_1 + b_{m2}v_2 + \cdots + b_{mn}v_n \end{aligned}$$

Probar que $\{u_i\}$ y $\{w_i\}$ generan el mismo espacio.

Efectuar una «operación elemental» de las del Problema 5.82 sobre $\{u_i\}$ equivale a efectuar una operación elemental entre filas sobre la matriz A. Como A y B son equivalentes por filas, B puede obtenerse de A por medio de una sucesión de operaciones elementales entre filas; por tanto, $\{w_i\}$ puede obtenerse de $\{u_i\}$ mediante la sucesión de operaciones correspondiente. De acuerdo con esto, $\{u_i\}$ y $\{w_i\}$ generan el mismo espacio.

5.84. Supóngase que $v_1, \ldots, v_n$ pertenecen a un espacio vectorial V sobre un cuerpo K. Sean

$$\begin{aligned} w_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ w_2 &= a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ w_n &= a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n \end{aligned}$$

donde $a_{ij} \in K$. Sea P la matriz n-cuadrada de los coeficientes, esto es, sea $P = (a_{ij})$.

a) Supóngase que P es invertible. Mostrar que $\{w_i\}$ y $\{v_i\}$ generan el mismo espacio y por ende $\{w_i\}$ es independiente si y sólo si lo es $\{v_i\}$.

b) Supóngase que P no es invertible. Mostrar que $\{w_i\}$ es dependiente.

c) Supóngase que $\{w_i\}$ es independiente. Mostrar que P es invertible.

a) Como P es invertible, es equivalente por filas a la matriz identidad I. Por tanto, según el problema precedente, $\{w_i\}$ y $\{v_i\}$ generan el mismo espacio. Así uno es independiente si y sólo si lo es el otro.

b) Dado que P no es invertible, es equivalente por filas a una matriz con una fila nula. Esto quiere decir que $\{w_i\}$ genera un espacio que tiene un conjunto generador con menos de n elementos. Siendo así, $\{w_i\}$ es dependiente.

c) Esta es la contrarrecíproca de la afirmación *b*), luego se deduce de *b*).

5.85. Supóngase que $A_1, A_2, \dots$ son conjuntos linealmente independientes de vectores y que $A_1 \subseteq A_2 \subseteq \cdots$. Probar que la unión $A = A_1 \cup A_2 \cup \cdots$ también es linealmente independiente.

Supongamos que A es linealmente dependiente. En ese caso existen vectores $v_1, \dots, v_n \in A$ y escalares $a_1, \dots, a_n \in K$, no todos 0, tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \qquad [1]$$

Como $A = \cup A_i$ y los $v_i \in A$, existen conjuntos $A_{i_1}, \dots, A_{i_n}$ tales que

$$v_1 \in A_{i_1},\ v_2 \in A_{i_2},\ \dots,\ v_n \in A_{i_n}$$

Sea k el máximo índice de los conjuntos $A_{i_j}: k = \text{máx}\,(i_1, \dots, i_n)$. Se sigue entonces, dado que $A_1 \subseteq A_2 \subseteq \cdots$, que cada A_{i_j} está contenido en A_k. Por consiguiente, $v_1, v_2, \dots, v_n \in A_k$ y así, de acuerdo con [1], A_k es linealmente dependiente, lo que contradice nuestra hipótesis. De este modo, A es linealmente independiente.

5.86. Sea K un subcuerpo de un cuerpo L que a su vez es subcuerpo de un cuerpo E; es decir, $K \subseteq L \subseteq E$. (Por tanto, K es subcuerpo de E.) Supóngase que E es de dimensión n sobre L y L de dimensión m sobre K. Demostrar que E es de dimensión mn sobre K.

Supongamos que $\{v_1, \dots, v_n\}$ es una base de E sobre L y $\{a_1, \dots, a_m\}$ una de L sobre K. Afirmamos que $\{a_i v_j : i = 1, \dots, m,\ j = 1, \dots, n\}$ es una base de E sobre K. Adviértase que $\{a_i v_j\}$ contiene mn elementos.

Sea w un elemento arbitrario de E. Como $\{v_1, \dots, v_n\}$ genera E sobre L, W es una combinación lineal de los v_i con coeficientes en L:

$$w = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n \qquad b_i \in L \qquad [1]$$

Como $\{a_1, \dots, a_m\}$ genera L sobre K, cada $b_i \in L$ es combinación lineal de los a_j con coeficientes en K:

$$\begin{aligned} b_1 &= k_{11} a_1 + k_{12} a_2 + \cdots + k_{1m} a_m \\ b_2 &= k_{21} a_1 + k_{22} a_2 + \cdots + k_{2m} a_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_n &= k_{n1} a_1 + k_{n2} a_2 + \cdots + k_{nm} a_m \end{aligned}$$

donde $k_{ij} \in K$. Sustituyendo en [1] obtenemos

$$\begin{aligned} w &= (k_{11} a_1 + \cdots + k_{1m} a_m) v_1 + (k_{21} a_1 + \cdots + k_{2m} a_m) v_2 + \cdots + (k_{n1} a_1 + \cdots + k_{nm} a_m) v_n = \\ &= k_{11} a_1 v_1 + \cdots + k_{1m} a_m v_1 + k_{21} a_1 v_2 + \cdots + k_{2m} a_m v_2 + \cdots + k_{n1} a_1 v_n + \cdots + k_{nm} a_m v_n = \\ &= \sum_{i,j} k_{ji} (a_i v_j) \end{aligned}$$

donde $k_{ji} \in K$. Siendo así, w es una combinación lineal de los $a_i v_j$ con coeficientes en K, por lo que $\{a_i v_j\}$ genera E sobre K.

La demostración será completa si probamos que $\{a_i v_j\}$ es linealmente independiente sobre K. Supongamos que, para escalares $x_{ji} \in K$, $\sum_{i,j} x_{ji}(a_i v_j) = 0$; esto es,

$$(x_{11} a_1 v_1 + x_{12} a_2 v_1 + \cdots + x_{1m} a_m v_1) + \cdots + (x_{n1} a_1 v_n + x_{n2} a_2 v_n + \cdots + x_{nm} a_m v_n) = 0$$

o $$(x_{11} a_1 + x_{12} a_2 + \cdots + x_{1m} a_m) v_1 + \cdots + (x_{n1} a_1 + x_{n2} a_2 + \cdots + x_{nm} a_m) v_n = 0$$

Dado que $\{v_1, \ldots, v_n\}$ es linealmente independiente sobre L y que los anteriores coeficientes de los v_i pertenecen a L, cada uno de los coeficientes debe ser 0:

$$x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m = 0, \ldots, x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m = 0$$

Pero $\{a_1, \ldots, a_m\}$ es linealmente independiente sobre K; por consiguiente, como los $x_{ji} \in K$,

$$x_{11} = 0, x_{12} = 0, \ldots, x_{1m} = 0, \ldots, x_{n1} = 0, x_{n2} = 0, \ldots, x_{nm} = 0$$

En consecuencia, $\{a_i v_j\}$ es linealmente independiente sobre K y queda demostrado el teorema.

PROBLEMAS SUPLEMENTARIOS

ESPACIOS VECTORIALES

5.87. Sea V el conjunto de los pares ordenados (a, b) de números reales con suma en V y producto por un escalar en V definidos según

$$(a, b) + (c, d) = (a + c, b + d) \qquad \text{y} \qquad k(a, b) = (ka, 0)$$

Demostrar que V satisface todos los axiomas de un espacio vectorial excepto $[M_4]$: $1u = u$. Por tanto, $[M_4]$ no es consecuencia de los otros axiomas.

5.88. Probar que el siguiente axioma $[A_4]$ puede derivarse de los otros axiomas de un espacio vectorial.

$[A_4]$ Para todo par de vectores $u, v \in V$, $u + v = v + u$.

5.89. Sea V el conjunto de sucesiones infinitas $(a_1, a_2, \ldots)$ en un cuerpo K con suma en V y producto por un escalar en V definidos por

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$
$$k(a_1, a_2, \ldots) = (ka_1, ka_2, \ldots)$$

donde a_i, b_j, $k \in K$. Probar que V es un espacio vectorial sobre K.

SUBESPACIOS

5.90. Determinar si W es o no un subespacio de $\mathbf{R}^3$, donde W consiste en aquellos vectores $(a, b, c) \in \mathbf{R}^3$ para los que *a*) $a = 2b$; *b*) $a \le b \le c$; *c*) $ab = 0$; *d*) $a = b = c$; *e*) $a = b^2$.

5.91. Sea V el espacio vectorial de las matrices n-cuadradas sobre un cuerpo K. Demostrar que W es un subespacio de V si W consiste en todas las matrices que son *a*) antisimétricas ($A^T = -A$), *b*) triangulares (superiores), *c*) diagonales, *d*) escalares.

5.92. Sea $AX = B$ un sistema inhomogéneo de ecuaciones lineales con n incógnitas sobre un cuerpo K. Demostrar que el conjunto solución del sistema no es un subespacio de K.

5.93. Discutir si $\mathbf{R}^2$ es o no subespacio de $\mathbf{R}^3$.

5.94. Supóngase que U y W son subespacios de V para los que $U \cup W$ es también subespacio. Probar que bien $U \subseteq W$ o bien $W \subseteq U$.

5.95. Sea V el espacio vectorial de todas las funciones del cuerpo real $\mathbf{R}$ en $\mathbf{R}$. Mostrar que W es un subespacio de V en cada uno de los siguientes casos.

a) W consiste en todas las funciones acotadas. [Aquí f: $\mathbf{R} \to \mathbf{R}$ es acotada si existe $M \in \mathbf{R}$ tal que $|f(x)| \leq M$, $\forall x \in \mathbf{R}$.]

b) W consiste en todas las funciones pares. [Aquí f: $\mathbf{R} \to \mathbf{R}$ es par si $f(-x) = f(x)$, $\forall x \in \mathbf{R}$.]

c) W consiste en todas las funciones continuas.

d) W consiste en todas las funciones diferenciables.

e) W consiste en todas las funciones integrables en, por ejemplo, el intervalo $0 \leq x \leq 1$.

(Los tres últimos casos requieren ciertos conocimientos de análisis.)

5.96. Sea V el espacio vectorial (Problema 5.89) de las sucesiones infinitas $(a_1, a_2, \ldots)$ en un cuerpo K. Demostrar que W es un subespacio de V, donde: *a*) W consiste en todas las sucesiones con 0 como primera componente, *b*) W consiste en todas las sucesiones con sólo un número finito de componentes no nulas.

COMBINACIONES LINEALES. ENVOLVENTES LINEALES

5.97. Demostrar que los números complejos $w = 2 + 3i$ y $z = 1 - 2i$ generan el cuerpo complejo $\mathbf{C}$ como espacio vectorial sobre el cuerpo real $\mathbf{R}$.

5.98. Probar que los polinomios $(1 - t)^3$, $(1 - t)^2$, $1 - t$ y 1 generan el espacio $\mathbf{P}_3(t)$ de los polinomios de grado ≤ 3.

5.99. Hallar un vector en $\mathbf{R}^3$ que genere la intersección de U y W, siendo U el plano xy: $U = \{(a, b, 0)\}$ y W el espacio generado por los vectores $(1, 2, 3)$ y $(1, -1, 1)$.

5.100. Demostrar que lin S es la intersección de todos los subespacios de V que contienen S.

5.101. Probar que lin S = lin $(S \cup \{0\})$. Esto es, añadir o suprimir el vector cero de un conjunto no altera el espacio generado por éste.

5.102. Probar que si $S \subseteq T$, necesariamente lin $S \subseteq$ lin T.

5.103. Probar que lin (lin S) = lin S.

5.104. Sean $W_1, W_2, \ldots$ subespacios de un espacio vectorial V tales que $W_1 \subseteq W_2 \subseteq \cdots$. Sea $W = W_1 \cup W_2 \cup \cdots$. *a*) Demostrar que W es un subespacio de V. *b*) Supóngase que S_i genera W_i para $i = 1, 2, \ldots$ Mostrar que $S = S_1 \cup S_2 \cup \cdots$ genera W.

DEPENDENCIA E INDEPENDENCIA LINEAL

5.105. Determinar si los siguientes vectores en $\mathbf{R}^4$ son linealmente dependientes o independientes:

a) $(1, 3, -1, 4), (3, 8, -5, 7), (2, 9, 4, 23)$ *b*) $(1, -2, 4, 1), (2, 1, 0, -3), (3, -6, 1, 4)$

5.106. Sea V el espacio vectorial de los polinomios de grado ≤ 3 sobre $\mathbf{R}$. Determinar si $u, v, w \in V$ son linealmente dependientes o independientes, donde:

a) $u = t^3 - 4t^2 + 2t + 3$, $v = t^3 + 2t^2 + 4t - 1$, $w = 2t^3 - t^2 - 3t + 5$.

b) $u = t^3 - 5t^2 - 2t + 3$, $v = t^3 - 4t^2 - 3t + 4$, $w = 2t^3 - 7t^2 - 7t + 9$.

5.107. Probar que *a)* los vectores $(1 - i, i)$ y $(2, -1 + i)$ en $\mathbf{C}^2$ son linealmente dependientes sobre el cuerpo complejo $\mathbf{C}$ pero son linealmente independientes sobre el cuerpo real $\mathbf{R}$; *b)* los vectores $(3 + \sqrt{2}, 1 + \sqrt{2})$ y $(7, 1 + 2\sqrt{2})$ en $\mathbf{R}^2$ son linealmente dependientes sobre el cuerpo real $\mathbf{R}$ pero son linealmente independientes sobre el cuerpo racional $\mathbf{Q}$.

5.108. Supóngase que $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ es un subconjunto linealmente independiente de V. Demostrar que $\text{lin}\, u_i \cap \text{lin}\, w_j = \{0\}$. (Recuérdese que lin u_i es el subespacio de V generado por los u_i.)

5.109. Supóngase que $v_1, v_2, \ldots, v_n$ son vectores linealmente independientes. Demostrar las siguientes afirmaciones:

a) $\{a_1 v_1, a_2 v_2, \ldots, a_n v_n\}$, donde cada $a_i \neq 0$ es linealmente independiente.

b) $\{v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n\}$, donde $w = b_1 v_1 + \cdots + b_i v_i + \cdots + b_n v_n$ y $b_i \neq 0$ es linealmente independiente.

5.110. Supóngase que $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$ son vectores linealmente independientes en K^n y que $v_1, \ldots, v_n$ son vectores linealmente independientes en un espacio vectorial V sobre K. Probar que los vectores

$$w_1 = a_{11}v_1 + \cdots + a_{1n}v_n, \ldots, w_m = a_{m1}v_1 + \cdots + a_{mn}v_n$$

también son linealmente independientes.

5.111. Supóngase que A es cualquier matriz n-cuadrada y que $u_1, u_2, \ldots, u_r$ son vectores columna $n \times 1$. Probar que si $Au_1, Au_2, \ldots, Au_r$ son vectores (columna) linealmente independientes, entonces $u_1, u_2, \ldots, u_r$ son linealmente independientes.

BASES Y DIMENSION

5.112. Encontrar un subconjunto de u_1, u_2, u_3, u_4 que sea base de $W = \text{lin}(u_1, u_2, u_3, u_4)$ de $\mathbf{R}^5$, donde:

a) $u_1 = (1, 1, 1, 2, 3)$, $u_2 = (1, 2, -1, -2, 1)$, $u_3 = (3, 5, -1, -2, 5)$, $u_4 = (1, 2, 1, -1, 4)$.

b) $u_1 = (1, -2, 1, 3, -1)$, $u_2 = (-2, 4, -2, -6, 2)$, $u_3 = (1, -3, 1, 2, 1)$, $u_4 = (3, -7, 3, 8, -1)$.

c) $u_1 = (1, 0, 1, 0, 1)$, $u_2 = (1, 1, 2, 1, 0)$, $u_3 = (1, 2, 3, 1, 1)$, $u_4 = (1, 2, 1, 1, 1)$.

d) $u_1 = (1, 0, 1, 1, 1)$, $u_2 = (2, 1, 2, 0, 1)$, $u_3 = (1, 1, 2, 3, 4)$, $u_4 = (4, 2, 5, 4, 6)$.

5.113. Sean U y W los siguientes subespacios de R^4:

$$U = \{(a, b, c, d) : b - 2c + d = 0\} \qquad W = \{(a, b, c, d) : a = d, b = 2c\}$$

Hallar una base y la dimensión de *a)* U, *b)* W, *c)* $U \cap W$.

5.114. Hallar una base y la dimensión del espacio solución W de cada uno de los sistemas homogéneos:

$$\begin{aligned} x + 2y - 2z + 2s - t &= 0 \\ x + 2y - z + 3s - 2t &= 0 \\ 2x + 4y - 7z + s + t &= 0 \end{aligned} \qquad \begin{aligned} x + 2y - z + 3s - 4t &= 0 \\ 2x + 4y - 2z - s + 5t &= 0 \\ 2x + 4y - 2z + 4s - 2t &= 0 \end{aligned}$$

a) *b)*

5.115. Encontrar un sistema homogéneo cuyo conjunto solución W esté generado por los tres vectores:

$$(1, -2, 0, 3, -1) \qquad (2, -3, 2, 5, -3) \qquad (1, -2, 1, 2, -2)$$

5.116. Sea V el espacio vectorial de los polinomios en t de grado $\leq n$. Determinar si cada uno de los siguientes conjuntos es o no una base de V:

a) $\{1, 1+t, 1+t+t^2, 1+t+t^2+t^3, \ldots, 1+t+t^2+\cdots+t^{n-1}+t^n\}$.

b) $\{1+t, t+t^2, t^2+t^3, \ldots, t^{n-2}+t^{n-1}, t^{n-1}+t^n\}$.

5.117. Hallar una base y la dimensión del subespacio W de $\mathbf{P}(t)$ generado por los polinomios

a) $u = t^3 + 2t^2 - 2t + 1$, $v = t^3 + 3t^2 - t + 4$ y $w = 2t^3 + t^2 - 7t - 7$.

b) $u = t^3 + t^2 - 3t + 2$, $v = 2t^3 + t^2 + t - 4$ y $w = 4t^3 + 3t^2 - 5t + 2$.

5.118. Sea V el espacio de las matrices 2×2 sobre $\mathbf{R}$. Encontrar una base y la dimensión del subespacio W de V generado por las matrices

$$\begin{pmatrix} 1 & -5 \\ -4 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ -1 & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & -4 \\ -5 & 7 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & -7 \\ -5 & 1 \end{pmatrix}$$

ESPACIO FILA Y RANGO DE UNA MATRIZ

5.119. Considérense los siguientes subespacios de $\mathbf{R}^3$:

$$U_1 = \text{lin}\,[(1, 1, -1), (2, 3, -1), (3, 1, -5)]$$
$$U_2 = \text{lin}\,[(1, -1, -3), (3, -2, -8), (2, 1, -3)]$$
$$U_3 = \text{lin}\,[(1, 1, 1), (1, -1, 3), (3, -1, 7)]$$

Determinar los que son idénticos entre sí.

5.120. Calcular el rango de cada matriz:

$$\begin{pmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & -3 & -2 & -3 \\ 1 & 3 & -2 & 0 & -4 \\ 3 & 8 & -7 & -2 & -11 \\ 2 & 1 & -9 & -10 & -3 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 \\ 4 & 5 & 5 \\ 5 & 8 & 1 \\ -1 & -2 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 3 & -7 \\ -6 & 1 \\ 5 & -8 \end{pmatrix}$$

a) *b)* *c)* *d)*

5.121. Demostrar que si se suprime una fila cualquiera de una matriz en forma escalonada (canónica por filas), la matriz resultante sigue estando en forma escalonada (canónica por filas).

5.122. Sean A y B matrices $m \times n$ arbitrarias. Probar que: rango $(A + B) \leq$ rango A + rango B.

5.123. Dar ejemplos de matrices 2×2 A y B tales que:

a) rango $(A + B) <$ rango A, rango B

b) rango $(A + B) =$ rango $A =$ rango B

c) rango $(A + B) >$ rango A, rango B

SUMAS. SUMAS DIRECTAS. INTERSECCIONES

5.124. Supóngase que U y W son subespacios bidimensionales de $\mathbf{R}^3$. Demostrar que $U \cap W \neq \{0\}$.

5.125. Supóngase que U y W son subespacios de V y que $\dim U = 4$, $\dim W = 5$ y $\dim V = 7$. Encontrar todas las dimensiones posibles de $U \cap W$.

5.126. Sean U y W subespacios de $\mathbf{R}^3$ para los que $\dim U = 1$, $\dim W = 2$ y $U \not\subseteq W$. Probar que $\mathbf{R}^3 = U \oplus W$.

5.127. Sean U el subespacio de $\mathbf{R}^5$ generado por

$$\{(1, 3, -3, -1, -4) \quad (1, 4, -1, -2, -2) \quad (2, 9, 0, -5, -2)\}$$

y W el generado por

$$\{(1, 6, 2, -2, 3) \quad (2, 8, -1, -6, -5) \quad (1, 3, -1, -5, -6)\}$$

Hallar: *a*) $\dim(U + W)$, *b*) $\dim(U \cap W)$.

5.128. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$. Hallar: *a*) $\dim(U + W)$, *b*) $\dim(U \cap W)$, donde

$$U = \text{lin}\,(t^3 + 4t^2 - t + 3,\ t^3 + 5t^2 + 5,\ 3t^3 + 10t^2 - 5t + 5)$$
$$W = \text{lin}\,(t^3 + 4t^2 + 6,\ t^3 + 2t^2 - t + 5,\ 2t^3 + 2t^2 - 3t + 9)$$

5.129. Sean U el subespacio de $\mathbf{R}^5$ generado por

$$(1, -1, -1, -2, 0) \qquad (1, -2, -2, 0, -3) \qquad \text{y} \qquad (1, -1, -2, -2, 1)$$

y W el generado por

$$(1, -2, -3, 0, -2) \qquad (1, -1, -3, 2, -4) \qquad \text{y} \qquad (1, -1, -2, 2, -5)$$

a) Encontrar dos sistemas homogéneos cuyos espacios solución sean U y W, respectivamente.

b) Encontrar una base y la dimensión de $U \cap W$.

5.130. Sean U_1, U_2 y U_3 los siguientes subespacios de $\mathbf{R}^3$:

$$U_1 = \{(a, b, c): a + b + c = 0\} \qquad U_2 = \{(a, b, c): a = c\} \qquad U_3 = \{(0, 0, c): c \in \mathbf{R}\}$$

Probar que: *a*) $\mathbf{R}^3 = U_1 + U_2$, *b*) $\mathbf{R}^3 = U_2 + U_3$, *c*) $\mathbf{R}^3 = U_1 + U_3$. ¿Cuándo es directa la suma?

5.131. Supóngase que U, V y W son subespacios de un espacio vectorial. Demostrar que

$$(U \cap V) + (U \cap W) \subseteq U \cap (V + W)$$

Encontrar subespacios de $\mathbf{R}^2$ para los que la igualdad no sea cierta.

5.132. La suma de dos subconjuntos (no necesariamente subespacios) no vacíos arbitrarios S y T de un espacio vectorial V se define como $S + T = \{s + t : s \in S, t \in T\}$. Probar que esta operación satisface:

a) Ley conmutativa: $S + T = T + S$

b) Ley asociativa: $(S_1 + S_2) + S_3 = S_1 + (S_2 + S_3)$

c) $S + \{0\} = \{0\} + S = S$

d) $S + V = V + S = V$

5.133. Supóngase que $W_1, W_2, \ldots, W_r$ son subespacios de un espacio vectorial V. Probar que:

a) $\text{lin}\,(W_1, W_2, \ldots, W_r) = W_1 + W_2 + \cdots + W_r$.

b) Si S_i genera W_i para $i = 1, \ldots, r$, entonces $S_1 \cup S_2 \cup \cdots \cup S_r$ genera $W_1 + W_2 + \cdots + W_r$.

5.134. Demostrar el Teorema 5.24.

5.135. Demostrar el Teorema 5.25.

5.136. Sean U y W espacios vectoriales sobre un cuerpo K. Sea V el conjunto de pares ordenados (u, w), donde u pertenece a U y w a W: $V = \{(u, w): u \in U, w \in W\}$. Probar que V es un espacio vectorial sobre K con suma en V y producto por un escalar en V definidos según

$$(u, w) + (u', w') = (u + u', w + w') \qquad \text{y} \qquad k(u, w) = (ku, kw)$$

donde $u, u' \in U$, $w, w' \in W$ y $k \in K$. (Este espacio V se denomina la *suma directa externa* de U y W.)

5.137. Sea V la suma directa externa de los espacios vectoriales U y W sobre un cuerpo K. (Véase el Problema 5.136.) Sean $\hat{U} = \{(u, 0): u \in U\}$, $\hat{W} = \{(0, w): w \in W\}$. Probar que

a) $\hat{U}$ y $\hat{W}$ son subespacios de V y que $V = \hat{U} \oplus \hat{W}$.

b) U es isomorfo a $\hat{U}$ bajo la correspondencia $u \leftrightarrow (u, 0)$ y que W es isomorfo a $\hat{W}$ bajo la correspondencia $w \leftrightarrow (0, w)$.

c) $\dim V = \dim U + \dim W$.

5.138. Supóngase $V = U \oplus W$. Sea $\hat{V}$ la suma directa externa de U y W. Probar que V es isomorfo a $\hat{V}$ bajo la correspondencia $v = u + w \leftrightarrow (u, w)$.

VECTORES COORDENADOS

5.139. Considérese la base $S = \{u_1 = (1, -2), u_2 = (4, -7)\}$ de $\mathbf{R}^2$. Hallar el vector coordenado $[v]$ de v relativo a S, siendo *a)* $v = (3, 5)$, *b)* $v = (1, 1)$, *c)* $v = (a, b)$.

5.140. Considérense el espacio vectorial $\mathbf{P}_3(t)$ de los polinomios de grado ≤ 3 y la base $S = \{1, t + 1, t^2 + t, t^3 + t^2\}$ de $\mathbf{P}_3(t)$. Encontrar el vector coordenado de v respecto a S, donde: *a)* $v = 2 - 3t + t^2 + 2t^3$; *b)* $v = 3 - 2t - t^2$; *c)* $v = a + bt + ct^2 + dt^3$.

5.141. Sea S la siguiente base del espacio vectorial W de las matrices 2×2 reales simétricas:

$$\left\{\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}\right\}$$

Encontrar el vector coordenado de la matriz $A \in W$ relativo a la base anterior S, donde: *a)* $A = \begin{pmatrix} 1 & -5 \\ -5 & 5 \end{pmatrix}$, *b)* $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.

CAMBIO DE BASE

5.142. Calcular la matriz de cambio de base P desde la base usual $E = \{(1, 0), (0, 1)\}$ de $\mathbf{R}^2$ hasta la base S, la matriz de cambio de base Q desde la base S regresando hasta la base E y el vector coordenado de $v = (a, b)$ relativo a S, donde

a) $S = \{(1, 2), (3, 5)\}$ *c)* $S = \{(2, 5), (3, 7)\}$

b) $S = \{(1, -3), (3, -8)\}$ *d)* $S = \{(2, 3), (4, 5)\}$

5.143. Considérense las siguientes bases de $\mathbf{R}^2$: $S = \{u_1 = (1, 2), u_2 = (2, 3)\}$ y $S' = \{v_1 = (1, 3), v_2 = (1, 4)\}$. Hallar: *a*) la matriz de cambio de base P de S a S', *b*) la matriz de cambio de base Q desde S' volviendo a S.

5.144. Supóngase que los ejes x e y en el plano $\mathbf{R}^2$ se giran 30° en el sentido contrario al de las agujas de un reloj proporcionando unos nuevos ejes x' e y' para el plano. Hallar: *a*) los vectores unitarios en las direcciones de los nuevos ejes, *b*) la matriz de cambio de base P para el nuevo sistema de coordenadas, *c*) las nuevas coordenadas de cada uno de los siguientes puntos respecto al nuevo sistema: $A(1, 3)$, $B(2, -5)$, $C(a, b)$.

5.145. Calcular la matriz de cambio de base P desde la base usual E de $\mathbf{R}^3$ hasta la base S, la matriz de cambio de base Q desde S hasta E y el vector coordenado de $v = (a, b, c)$ relativo a S, estando S constituida por los vectores:

a) $u_1 = (1, 1, 0), u_2 = (0, 1, 2), u_3 = (0, 1, 1)$

b) $u_1 = (1, 0, 1), u_2 = (1, 1, 2), u_3 = (1, 2, 4)$

c) $u_1 = (1, 2, 1), u_2 = (1, 3, 4), u_3 = (2, 5, 6)$

5.146. Supóngase que S_1, S_2 y S_3 son bases de un espacio vectorial V y que P y Q son las matrices de cambio de base desde S_1 hasta S_2 y desde S_2 hasta S_3, respectivamente. Demostrar que el producto PQ es la matriz de cambio de base desde S_1 hasta S_3.

PROBLEMAS VARIOS

5.147. Determinar la dimensión del espacio vectorial W de las matrices n-cuadradas sobre un cuerpo K: *a*) simétricas, *b*) antisimétricas.

5.148. Sean V un espacio vectorial de dimensión n sobre un cuerpo K y K uno de dimensión m sobre un subcuerpo F. (Por tanto, V puede verse también como un espacio vectorial sobre el subcuerpo F.) Demostrar que la dimensión de V sobre F es mn.

5.149. Sean $t_1, t_2, \ldots, t_n$ símbolos y K cualquier cuerpo. Sea V el conjunto de expresiones

$$a_1t_1 + a_2t_2 + \cdots + a_nt_n \qquad \text{donde} \qquad a_i \in K$$

Defínase la suma en V según

$$(a_1t_1 + a_2t_2 + \cdots + a_nt_n) + (b_1t_1 + b_2t_2 + \cdots + b_nt_n) = (a_1 + b_1)t_1 + (a_2 + b_2)t_2 + \cdots + (a_n + b_n)t_n$$

Defínase el producto por un escalar en V según

$$k(a_1t_1 + a_2t_2 + \cdots + a_nt_n) = ka_1t_1 + ka_2t_2 + \cdots + ka_nt_n$$

Probar que V es un espacio vectorial sobre K con las operaciones anteriores. Probar, asimismo, que $\{t_1, \ldots, t_n\}$ es una base de V, donde, para $i = 1, \ldots, n$,

$$t_i = 0t_1 + \cdots + 0t_{i-1} + 1t_i + 0t_{i+1} + \cdots + 0t_n$$

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

5.90. *a*) Sí.

b) No; por ejemplo, $(1, 2, 3) \in W$ pero $-2(1, 2, 3) \notin W$.

c) No; por ejemplo, $(1, 0, 0), (0, 1, 0) \in W$, pero no su suma.

d) Sí.

e) No; por ejemplo, $(9, 3, 0) \in W$ pero $2(9, 3, 0) \notin W$.

5.92. $X = 0$ no es solución de $AX = B$.

5.93. No. Aunque uno pueda «identificar» el vector $(a, b) \in \mathbf{R}^2$ con, digamos $(a, b, 0)$ en el plano xy de $\mathbf{R}^3$, son elementos distintos pertenecientes a conjuntos distintos, disjuntos.

5.95. *a*) Sean $f, g \in W$ con cotas M_f y M_g, respectivamente. En tal caso, para todo par de escalares $a, b \in \mathbf{R}$,

$$|(af + bg)(x)| = |af(x) + bg(x)| \leq |af(x)| + |bg(x)| = |a||f(x)| + |b||g(x)| \leq |a|M_f + |b|M_g$$

Esto es, $|a| M_f + |b| M_g$ es una cota para la función $af + bg$.

b) $(af + bg)(-x) = af(-x) + bg(-x) = af(x) + bg(x) = (af + bg)(x)$.

5.99. $(2, -5, 0)$.

5.105. *a*) Dependientes, *b*) Independientes.

5.106. *a*) Independientes, *b*) Dependientes.

5.107. *a*) $(2, -1 + i) = (1 + i)(1 - i, i)$; *b*) $(7, 1 + 2\sqrt{2}) = (3 - \sqrt{2})(3 + \sqrt{2}, 1 + \sqrt{2})$.

5.112. *a*) u_1, u_2, u_4; *b*) u_1, u_3; *c*) u_1, u_2, u_3, u_4; *d*) u_1, u_2, u_3.

5.113. *a*) Base, $\{(1, 0, 0, 0), (0, 2, 1, 0), (0, -1, 0, 1)\}$; $\dim U = 3$.

b) Base, $\{(1, 0, 0, 1), (0, 2, 1, 0)\}$; $\dim W = 2$.

c) Base, $\{(0, 2, 1, 0)\}$; $\dim (U \cap W) = 1$. *Indicación:* $U \cap W$ debe satisfacer las tres condiciones sobre a, b, c y d.

5.114. *a*) Base, $\{(2, -1, 0, 0, 0), (4, 0, 1, -1, 0), (3, 0, 1, 0, 1)\}$; $\dim W = 3$.

b) Base, $\{(2, -1, 0, 0, 0), (1, 0, 1, 0, 0)\}$; $\dim W = 2$.

5.115. $\begin{cases} 5x + y - z - s = 0 \\ x + y - z - t = 0 \end{cases}$

5.116. *a*) Sí. *b*) No, porque $\dim V = n + 1$, pero el conjunto contiene sólo n elementos.

5.117. *a*) $\dim W = 2$, *b*) $\dim W = 3$.

5.118. $\dim W = 2$.

5.119. U_1 y U_2.

5.120. *a*) 3, *b*) 2, *c*) 3, *d*) 2.

5.123. *a*) $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}$ *c*) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

b) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$

5.125. $\dim (U \cap W) = 2$, 3 ó 4.

5.127. *a*) $\dim (U + W) = 3$, *b*) $\dim (U \cap W) = 2$.

5.128. *a*) $\dim (U + W) = 3$, $\dim (U \cap W) = 1$.

5.129. *a*) $\begin{cases} 3x + 4y - z \quad\quad - t = 0 \\ 4x + 2y \quad\quad + s \quad\quad = 0 \end{cases}$, $\begin{cases} 4x + 2y \quad\quad - s \quad\quad = 0 \\ 9x + 2y + z \quad\quad + t = 0 \end{cases}$

b) $\{(1, -2, -5, 0, 0), (0, 0, 1, 0, -1)\}$. $\dim (U \cap W) = 2$.

5.130. La suma es directa en *b*) y *c*).

5.131. En $\mathbf{R}^2$, sean U, V y W, respectivamente, la recta $y = x$ y los ejes x e y.

5.139. *a*) $[-41, 11]$, *b*) $[-11, 3]$, *c*) $[-7a - 4b, 2a + b]$.

5.140. *a*) $[4, -2, -1, 2]$, *b*) $[4, -1, -1, 0]$, *c*) $[a - b + c - d, b - c + d, c - d, d]$.

5.141. *a*) $[2, -1, 1]$, *b*) $[3, 1, -2]$.

5.142. *a*) $P = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$, $Q = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$, $[v] = \begin{pmatrix} -5a + b \\ 2a - b \end{pmatrix}$

b) $P = \begin{pmatrix} 1 & 3 \\ -3 & -8 \end{pmatrix}$, $Q = \begin{pmatrix} -8 & -3 \\ 3 & 1 \end{pmatrix}$, $[v] = \begin{pmatrix} -8a - 3b \\ 3a - b \end{pmatrix}$

c) $P = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$, $Q = \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}$, $[v] = \begin{pmatrix} -7a + 3b \\ 5a - 2b \end{pmatrix}$

d) $P = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$, $Q = \begin{pmatrix} -\frac{5}{2} & 2 \\ \frac{3}{2} & -1 \end{pmatrix}$, $[v] = \begin{pmatrix} (-\frac{5}{2})a + 2b \\ (\frac{3}{2})a - b \end{pmatrix}$

5.143. *a*) $P = \begin{pmatrix} 3 & 5 \\ -1 & -1 \end{pmatrix}$, *b*) $Q = \begin{pmatrix} 2 & 5 \\ -1 & -3 \end{pmatrix}$

5.144. *a*) $(\sqrt{3}/2, \frac{1}{2}), (-\frac{1}{2}, \sqrt{3}/2)$.

b) $P = \begin{pmatrix} \sqrt{3}/2 & -\frac{1}{2} \\ \frac{1}{2} & \sqrt{3}/2 \end{pmatrix}$

c) $[A] = [(\sqrt{3} - 3)/2, (1 + 3\sqrt{3})/2]$,
$[B] = [(2\sqrt{3} + 5)/2, (2 - 5\sqrt{3})/2]$,
$[C] = [(\sqrt{3}a - b)/2, (a + \sqrt{3}b)/2]$.

5.145. Como E es la base usual, sea P sencillamente la matriz cuyas columnas son u_1, u_2, u_3. En tal caso, $Q = P^{-1}$ y $[v] = P^{-1}v = Qv$.

a) $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ -2 & 2 & -1 \end{pmatrix}$, $[v] = \begin{pmatrix} a \\ a - b + c \\ -2a + 2b - c \end{pmatrix}$

b) $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}$, $[v] = \begin{pmatrix} -2b + c \\ 2a + 3b - 2c \\ -a - b + c \end{pmatrix}$

c) $P = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 5 \\ 1 & 4 & 6 \end{pmatrix}$, $Q = \begin{pmatrix} -2 & 2 & -1 \\ -7 & 4 & -1 \\ 5 & -3 & 1 \end{pmatrix}$, $[v] = \begin{pmatrix} -2a + 2b - c \\ -7a + 4b - c \\ 5a - 3b + c \end{pmatrix}$

5.147. *a*) $n(n + 1)/2$, *b*) $n(n - 1)/2$.

5.148. *Indicación:* La demostración es casi idéntica a la dada en el Problema 5.86 para el caso especial en el que V es una extensión del cuerpo K.

CAPITULO 6

Espacios con producto interno. Ortogonalidad

6.1. INTRODUCCION

La definición de un espacio vectorial V involucra un cuerpo arbitrario K. En este capítulo nos restringiremos a los casos en que K es el cuerpo real $\mathbf{R}$ o el complejo $\mathbf{C}$. En concreto, supondremos primero, a menos que se establezca o sobrentienda lo contrario, $K = \mathbf{R}$, en cuyo caso V se denomina *espacio vectorial real*, y en las últimas secciones extenderemos nuestros resultados al caso $K = \mathbf{C}$, en cuyo caso V recibe el nombre de *espacio vectorial complejo.*

Recordemos que los conceptos de «longitud» y «ortogonalidad» no aparecían en la investigación de los espacios vectoriales arbitrarios (aunque lo hicieran en el Capítulo 2, referente a los espacios $\mathbf{R}^n$ y $\mathbf{C}^n$). En este capítulo establecemos una estructura adicional en un espacio vectorial V para obtener un espacio con producto interno. Es en este contexto en el que se definen los conceptos citados.

Como en el Capítulo 5, adoptamos la siguiente notación (a menos que se especifique o venga implícita otra cosa):

V	el espacio vectorial dado
u, v, w	vectores en V
K	el cuerpo de escalares
a, b, c o k	escalares en K

Subrayamos que V denotará un espacio vectorial de dimensión finita salvo que se establezca o sobrentienda lo contrario. De hecho, muchos de los teoremas del capítulo no son válidos para espacios de dimensión infinita. Esto se ilustrará en algunos de los ejemplos y problemas.

6.2. ESPACIOS CON PRODUCTO INTERNO

Comenzamos con una definición.

Definición: Sea V un espacio vectorial real. Supongamos que a cada par de vectores $u, v \in V$ se le asigna un número real, denotado por $\langle u, v\rangle$. Esta función se llama un *producto interno* (*real*) en V si satisface los axiomas:

$[I_1]$ (Propiedad lineal) $\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$.

$[I_2]$ (Propiedad simétrica) $\langle u, v\rangle = \langle v, u\rangle$.

$[I_3]$ (Propiedad definida positiva) $\langle u, u\rangle \geqslant 0$; y $\langle u, u\rangle = 0$ si y sólo si $u = 0$.

El espacio vectorial V se denomina entonces *espacio* (*real*) *con producto interno*.

El axioma $[I_1]$ es equivalente a las dos condiciones:

$$a)\quad \langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle \qquad \text{y} \qquad b)\quad \langle ku, v\rangle = k\langle u, w\rangle$$

Usando $[I_1]$ y el axioma de simetría $[I_2]$ llegamos a

$$\langle u, cv_1 + dv_2\rangle = \langle cv_1 + dv_2, u\rangle = c\langle v_1, u\rangle + d\langle v_2, u\rangle = c\langle u, v_1\rangle + d\langle u, v_2\rangle$$

o, equivalentemente, a las dos condiciones

$$a)\quad \langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle \qquad \text{y} \qquad b)\quad \langle u, kv\rangle = k\langle u, v\rangle$$

Esto es, la función producto interno es también lineal en su segunda posición (variable). Por inducción tendremos

$$\langle a_1u_1 + \cdots + a_ru_r, v\rangle = a_1\langle u_1, v\rangle + a_2\langle u_2, v\rangle + \cdots + a_r\langle u_r, v\rangle$$

y

$$\langle u, b_1v_1 + b_2v_2 + \cdots + b_sv_s\rangle = b_1\langle u, v_1\rangle + b_2\langle u, v_2\rangle + \cdots + b_s\langle u, v_s\rangle$$

Combinar estas dos propiedades nos conduce a la fórmula general escrita a continuación:

$$\left\langle \sum_{i=1}^{r} a_iu_i, \sum_{j=1}^{s} b_jv_j \right\rangle = \sum_{i=1}^{r} \sum_{j=1}^{s} a_ib_j\langle u_i, v_j\rangle$$

Podemos hacer, por orden, las siguientes observaciones:

Nota 1: El axioma $[I_1]$ por sí mismo implica

$$\langle 0, 0\rangle = \langle 0v, 0\rangle = 0\langle v, 0\rangle = 0$$

En consecuencia, $[I_1]$, $[I_2]$ e $[I_3]$ son equivalentes a $[I_1]$, $[I_2]$ y el axioma:

$$[I'_3] \quad \text{Si } u \neq 0, \text{ necesariamente } \langle u, u\rangle > 0$$

O sea, una función que satisface $[I_1]$, $[I_2]$ e $[I'_3]$ es un producto interno.

Nota 2: De acuerdo con $[I_3]$, $\langle u, u\rangle$ es no negativo y por tanto existe su raíz cuadrada real positiva. Utilizamos la notación

$$\|u\| = \sqrt{\langle u, u\rangle}$$

El número real no negativo $\|u\|$ se denomina la *norma* o *longitud* de u. Esta función satisface los axiomas de una norma para un espacio vectorial. (Véanse el Teorema 6.25 y la Sección 6.9.) Hacemos notar que la relación $\|u\|^2 = \langle u, u \rangle$ se empleará con frecuencia.

EJEMPLO 6.1. Consideremos el espacio vectorial $\mathbf{R}^n$. El *producto escalar* en $\mathbf{R}^n$ se define según

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

donde $u = (a_i)$ y $v = (b_i)$. Esta función define un producto interno en $\mathbf{R}^n$. La norma $\|u\|$ del vector $u = (a_i)$ en este espacio es:

$$\| u \| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

Por otra parte, por el teorema de Pitágoras sabemos que la distancia del origen O en $\mathbf{R}^3$ al punto $P(a, b, c)$, mostrado en la Figura 6-1, viene dada por $\sqrt{a^2 + b^2 + c^2}$. Esta coincide precisamente con la norma del vector $v = (a, b, c)$ en $\mathbf{R}^3$ definida anteriormente. Dado que el teorema de Pitágoras es una consecuencia de los axiomas de la Geometría Euclídea, el espacio vectorial $\mathbf{R}^n$ con el producto interno y norma precedentes se conoce como *n-espacio euclídeo*. A pesar de haber muchos otros posibles, supondremos que este es el producto interno de $\mathbf{R}^n$, a no ser que se establezca o sobrentienda lo contrario. Se llama el *producto interno usual* de $\mathbf{R}^n$.

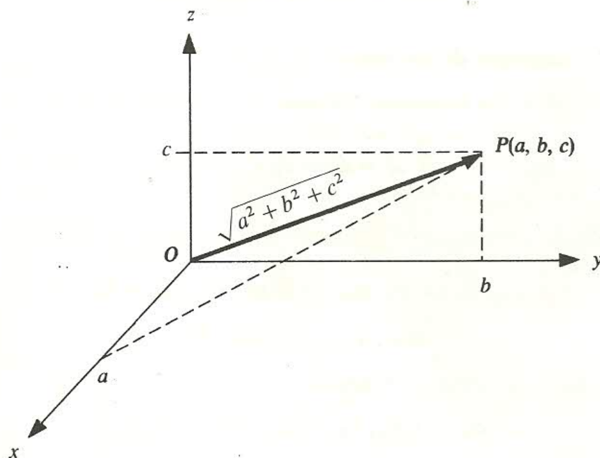


Figura 6-1.

Nota: Los vectores en $\mathbf{R}^n$ se representan frecuentemente por matrices columna $n \times 1$. En tal caso, el producto interno usual (Ejemplo 6.1) puede definirse según

$$\langle u, v \rangle = u^T v$$

EJEMPLO 6.2

a) Sea V el espacio vectorial de las funciones reales continuas en el intervalo $a \leqslant t \leqslant b$. El siguiente es un producto interno en V:

$$\langle f, g \rangle = \int_a^b f(t) g(t) \, dt$$

donde $f(t)$ y $g(t)$ son ahora funciones continuas cualesquiera en $[a, b]$.

b) Sea V nuevamente el espacio vectorial de las funciones reales continuas en el intervalo $a \leqslant t \leqslant b$. Si $w(t)$ es una función continua dada, positiva en $[a, b]$, otro producto interno en V es:

$$\langle f, g \rangle = \int_a^b w(t) f(t) g(t)\, dt$$

En este caso, $w(t)$ se denomina una *función peso* para el producto interno.

EJEMPLO 6.3

a) Denotemos por V el espacio vectorial de las matrices $m \times n$ sobre **R**. Un producto interno en V es:

$$\langle A, B \rangle = \operatorname{tr}(B^T A)$$

donde tr quiere decir traza o suma de los elementos diagonales. Si $A = (a_{ij})$ y $B = (b_{ij})$,

$$\langle A, B \rangle = \operatorname{tr}(B^T A) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}$$

que es la suma de los productos de las entradas correspondientes. En particular,

$$\| A \|^2 = \langle A, A \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2$$

que es la suma de los cuadrados de los elementos de A.

b) Sea V el espacio vectorial de las sucesiones infinitas de números reales $(a_1, a_2, \ldots)$ que cumplen

$$\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + \cdots < \infty$$

esto es, la suma converge. La suma y el producto por un escalar se definen por componentes:

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$
$$k(a_1, a_2, \ldots) = (ka_1, ka_2, \ldots)$$

Puede definirse un producto interno en V según

$$\langle (a_1, a_2, \ldots), (b_1, b_2, \ldots) \rangle = a_1 b_1 + a_2 b_2 + \cdots$$

La suma anterior converge absolutamente para todo par de vectores en V (Problema 6.12); por consiguiente, el producto interno está bien definido. Este espacio con producto interno recibe el nombre de *espacio* l_2 (o *espacio de Hilbert*).

6.3. DESIGUALDAD DE CAUCHY-SCHWARZ. APLICACIONES

La fórmula escrita abajo (demostrada en el Problema 6.10) se conoce como desigualdad de Cauchy-Schwarz, y se emplea en numerosas ramas de las matemáticas.

Teorema 6.1 (Cauchy-Schwarz): Para todo par de vectores $u, v \in V$,

$$\langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle \quad \text{o, equivalentemente,} \quad |\langle u, v \rangle| \leq \| u \| \, \| v \|$$

A continuación examinamos la desigualdad en casos específicos.

EJEMPLO 6.4

a) Consideremos reales arbitrarios $a_1, \ldots, a_n, b_1, \ldots, b_n$. Por la desigualdad de Cauchy-Schwarz,

$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

es decir, $(u \cdot v)^2 \leqslant \|u\|^2 \|v\|^2$, donde $u = (a_i)$ y $v = (b_i)$.

b) Sean f y g funciones reales continuas arbitrarias definidas en el intervalo unidad $0 \leqslant t \leqslant 1$. Según la desigualdad de Cauchy-Schwarz,

$$(\langle f, g \rangle)^2 = \left(\int_0^1 f(t)g(t)\,dt\right)^2 \leq \int_0^1 f^2(t)\,dt \int_0^1 g^2(t)\,dt = \|f\|^2 \|g\|^2$$

Aquí V es el espacio con producto interno del Ejemplo 6.2 *a*).

El teorema que sigue (demostrado en el Problema 6.11) nos da las propiedades básicas de una norma; la demostración de la tercera requiere el uso de la desigualdad de Cauchy-Schwarz.

Teorema 6.2: Sea V un espacio con producto interno. La norma en V satisface las propiedades:

$[N_1]$ $\|v\| \geqslant 0$; y $\|v\| = 0$ si y sólo si $v = 0$.

$[N_2]$ $\|kv\| = |k| \, \|v\|$.

$[N_3]$ $\|u + v\| \leqslant \|u\| + \|v\|$.

Las propiedades $[N_1]$, $[N_2]$ y $[N_3]$ precedentes son las que se han elegido como axiomas de una norma abstracta en un espacio vectorial (véase la Sección 6.9). El teorema anterior dice, pues, que la definida por un producto interno es realmente una norma. La propiedad $[N_3]$ se llama a menudo *desigualdad triangular* porque, si vemos $u + v$ como el lado del triángulo formado con u y v (como se muestra en la Figura 6-2), $[N_3]$ establece que la longitud de un lado de un triángulo es menor o igual que la suma de las longitudes de los otros dos.

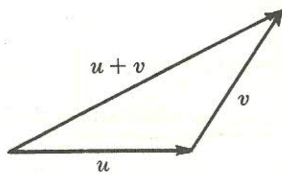


Figura 6-2.

Podemos hacer, por orden, una serie de observaciones.

Nota 1: Si $\|u\| = 1$, o, equivalentemente, si $\langle u, u \rangle = 1$, u se denomina un *vector unitario* y se dice que está *normalizado*. Todo vector no nulo $v \in V$ puede multiplicarse por el inverso de su longitud para obtener el vector unitario

$$\hat{v} = \frac{1}{\|v\|} v$$

que es un múltiplo positivo de v. Este proceso se conoce como *normalización* de v.

Nota 2: El número real no negativo $d(u, v) = \|u - v\|$ recibe el nombre de distancia entre u y v; esta función satisface los axiomas de un espacio métrico (véase el Teorema 6.19).

Nota 3: Para todo par de vectores no nulos $u, v \in V$, el ángulo entre u y v se define como el ángulo θ tal que $0 \leqslant \theta \leqslant \pi$ y

$$\cos \theta = \frac{\langle u, v \rangle}{\| u \| \, \| v \|}$$

Según la desigualdad de Cauchy-Schwarz, $-1 \leqslant \cos \theta \leqslant 1$ y así el ángulo θ siempre existe y es único.

6.4. ORTOGONALIDAD

Sea V un espacio con producto interno. Se dice que los vectores $u, v \in V$ son *ortogonales* y que u es *ortogonal* a v si

$$\langle u, v \rangle = 0$$

La relación es claramente simétrica; es decir, si u es ortogonal a v, necesariamente $\langle v, u \rangle = 0$ y por tanto v es ortogonal a u. Hacemos notar que $0 \in V$ es ortogonal a todo $v \in V$, ya que

$$\langle 0, v \rangle = \langle 0v, v \rangle = 0\langle v, v \rangle = 0$$

Recíprocamente, si u es ortogonal a todo $v \in V$, entonces $\langle u, u \rangle = 0$, luego $u = 0$ por $[I_3]$. Observemos que u y v son ortogonales si y sólo si $\cos \theta = 0$, siendo θ el ángulo entre u y v, lo que es cierto si y sólo si u y v son «perpendiculares», esto es, si $\theta = \pi/2$ (o $\theta = 90°$).

EJEMPLO 6.5

a) Consideremos un vector arbitrario $u = (a_1, a_2, \ldots, a_n)$ en $\mathbf{R}^n$. Un vector $v = (x_1, x_2, \ldots, x_n)$ es ortogonal a u si

$$\langle u, v \rangle = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

Dicho de otro modo, v es ortogonal a u si satisface una ecuación homogénea en la que los coeficientes son los elementos de u.

b) Supongamos que buscamos un vector no nulo que sea ortogonal a $v_1 = (1, 3, 5)$ y $v_2 = (0, 1, 4)$ en $\mathbf{R}^3$. Sea $w = (x, y, z)$. Queremos que

$$0 = \langle v_1, w \rangle = x + 3y + 5z \qquad \text{y} \qquad 0 = \langle v_2, w \rangle = y + 4z$$

Así obtenemos el sistema homogéneo

$$x + 3y + 5z = 0 \qquad y + 4z = 0$$

Tomamos $z = 1$ llegando a $y = -4$ y $x = 7$; por tanto, $w = (7, -4, 1)$ es ortogonal a v_1 y v_2. Normalizando w conseguimos

$$\hat{w} = w/\| w \| = (7/\sqrt{66}, -4/\sqrt{66}, 1/\sqrt{66})$$

que es un vector unitario ortogonal a v_1 y v_2.

COMPLEMENTOS ORTOGONALES

Sea S un subconjunto de un espacio con producto interno V. El complemento ortogonal de S, denotado por $S^\perp$ (leído «S perp»), consiste en aquellos vectores de V que son ortogonales a todo vector $u \in S$:

$$S^\perp = \{v \in V : \langle v, u \rangle = 0 \text{ para todo } u \in S\}$$

En particular, para un vector dado u en V, tendremos

$$u^\perp = \{v \in V : \langle v, u \rangle = 0\}$$

Es decir, $u^\perp$ consiste en todos los vectores que son ortogonales al vector dado u.

Probemos que $S^\perp$ es un subespacio de V. Obviamente, $0 \in S^\perp$, puesto que 0 es ortogonal a todo vector en V. Supongamos ahora que $v, w \in S^\perp$. En tal caso, para todo par de escalares a y b y todo vector $u \in S$,

$$\langle av + bw, u \rangle = a\langle v, u \rangle + b\langle w, u \rangle = a \cdot 0 + b \cdot 0 = 0$$

De este modo, $av + bw \in S^\perp$ y $S^\perp$ es un subespacio de V.

Establezcamos formalmente este resultado.

Proposición 6.3: Sea S un subconjunto de un espacio con producto interno V. Entonces $S^\perp$ es un subespacio de V.

Nota 1: Supongamos que u es un vector no nulo en $\mathbf{R}^3$. En este caso existe una interpretación geométrica de $u^\perp$. De forma específica, $u^\perp$ es el plano en $\mathbf{R}^3$ que, pasando por el origen, es perpendicular al vector u, tal y como se muestra en la Figura 6-3.

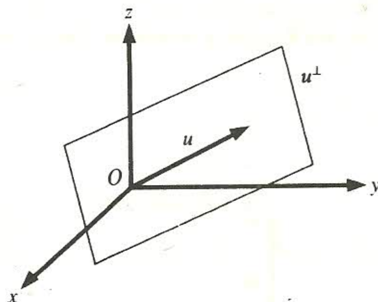


Figura 6-3.

Nota 2: Consideremos un sistema de ecuaciones lineales homogéneo sobre $\mathbf{R}$:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}$$

Recordemos que el espacio solución W puede verse como la solución de la ecuación matricial equivalente $AX = 0$, donde $A = (a_{ij})$ y $X = (x_i)$. Esto suministra otra interpretación de W, en la que se usa la noción de ortogonalidad. Concretamente, cada vector solución $v = (x_1, x_2, \ldots, x_n)$ es ortogonal a todas las filas de A; y, en consecuencia, W es el complemento ortogonal del espacio fila de A.

EJEMPLO 6.6. Supongamos que queremos hallar una base del subespacio $u^\perp$ en $\mathbf{R}^3$, siendo $u = (1, 3, -4)$. Nótese que $u^\perp$ está formado por todos los vectores (x, y, z) tales que

$$\langle (x, y, z), (1, 3, -4) \rangle = 0 \qquad \text{o} \qquad x + 3y - 4z = 0$$

Las variables libres son y y z. Tomemos:

1. $y = -1$, $z = 0$ para obtener la solución $w_1 = (3, -1, 0)$.
2. $y = 0$, $z = 1$ para obtener la solución $w_2 = (4, 0, 1)$.

Los vectores w_1 y w_2 constituyen una base del espacio solución de la ecuación y por ende una base de $u^\perp$.

Supongamos que W es un subespacio de V. W y $W^\perp$ son ambos subespacios de V. El teorema enunciado a continuación, cuya demostración (Problema 6.35) requiere el uso de resultados de las secciones posteriores, es básico dentro del álgebra lineal.

Teorema 6.4: Sea W un subespacio de V. Entonces V es la suma directa de W y $W^\perp$, o sea, $V = W \oplus W^\perp$.

EJEMPLO 6.7. Sea W el eje z en $\mathbf{R}^3$, esto es, $W = \{(0, 0, c) : c \in \mathbf{R}\}$. En ese caso, $W^\perp$ es el plano xy o, en otras palabras, $W^\perp = \{(a, b, 0) : a, b \in \mathbf{R}\}$, tal y como se ilustra en la Figura 6-4. Como se señaló previamente, $\mathbf{R}^3 = W \oplus W^\perp$.

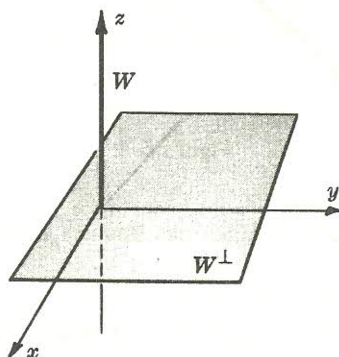


Figura 6-4.

6.5. CONJUNTOS ORTOGONALES Y BASES. PROYECCIONES

Un conjunto de vectores S en V se dice *ortogonal* si cada par de vectores en S lo son, y S se dice *ortonormal* si es ortogonal y cada vector de S tiene longitud unidad. Dicho de otro modo, $S = \{u_1, u_2, \ldots, u_r\}$ es *ortogonal* si

$$\langle u_i, u_j \rangle = 0 \quad \text{para} \quad i \neq j$$

y *ortonormal* si

$$\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{para } i \neq j \\ 1 & \text{para } i = j \end{cases}$$

La expresión *normalizar* un conjunto ortogonal S de vectores se refiere al proceso de multiplicar cada vector de S por el inverso de su longitud con el fin de transformar S en un conjunto ortonormal de vectores.

Una base S de un espacio vectorial V se llama *base ortogonal* u *ortonormal* según lo sea como conjunto de vectores.

Son aplicables los siguientes teoremas, demostrados en los Problemas 6.20 y 6.21, respectivamente.

Teorema 6.5: Supongamos que S es un conjunto ortogonal de vectores no nulos. En tal caso, S es linealmente independiente.

Teorema 6.6 (Pitágoras): Supongamos que $\{u_1, u_2, \ldots, u_r\}$ es un conjunto ortogonal de vectores. Entonces

$$\| u_1 + u_2 + \cdots + u_r \|^2 = \| u_1 \|^2 + \| u_2 \|^2 + \cdots + \| u_r \|^2$$

Probemos aquí el teorema de Pitágoras en el caso especial y familiar de dos vectores. Específicamente, supongamos $\langle u, v \rangle = 0$. Tenemos

$$\| u + v \|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle v, v \rangle = \| u \|^2 + \| v \|^2$$

lo que nos lleva a nuestro resultado.

EJEMPLO 6.8

a) Consideremos la base usual E del 3-espacio euclídeo $\mathbf{R}^3$:

$$E = \{e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)\}$$

Es claro que

$$\langle e_1, e_2 \rangle = \langle e_1, e_3 \rangle = \langle e_2, e_3 \rangle = 0 \qquad \text{y} \qquad \langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = \langle e_3, e_3 \rangle = 1$$

Así pues, E es una base ortonormal de $\mathbf{R}^3$. Con mayor generalidad, la base usual de $\mathbf{R}^n$ es ortonormal para todo n.

b) Sea V el espacio vectorial de las funciones reales continuas sobre el intervalo $-\pi \leqslant t \leqslant \pi$ con producto interno definido por $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t)dt$. Un ejemplo clásico de subconjunto ortogonal de V es:

$$\{1, \cos t, \cos 2t, \ldots, \operatorname{sen} t, \operatorname{sen} 2t, \ldots\}$$

El conjunto ortogonal precedente juega un papel fundamental en la teoría de las series de Fourier.

c) Consideremos el conjunto S de vectores en $\mathbf{R}^4$:

$$S = \{u = (1, 2, -3, 4), v = (3, 4, 1, -2), w = (3, -2, 1, 1\}$$

Nótese que

$$\langle u, v\rangle = 3 + 8 - 3 + 8 = 0 \qquad \langle u, w\rangle = 3 - 4 - 3 + 4 = 0 \qquad \langle v, w\rangle = 9 - 8 + 1 - 2 = 0$$

De este modo, S es ortogonal. Normalizamos S para conseguir un conjunto ortonormal, calculando primero

$$\|u\|^2 = 1 + 4 + 9 + 16 = 30 \qquad \|v\|^2 = 9 + 16 + 1 + 4 = 30 \qquad \|w\|^2 = 9 + 4 + 1 + 1 = 15$$

El que sigue es, pues, el conjunto ortonormal de vectores deseado:

$$\hat{u} = (1/\sqrt{30}, 2/\sqrt{30}, -3/\sqrt{30}, 4/\sqrt{30})$$
$$\hat{v} = (3/\sqrt{30}, 4/\sqrt{30}, 1/\sqrt{30}, -2/\sqrt{30})$$
$$\hat{w} = (3/\sqrt{15}, -2/\sqrt{30}, 1/\sqrt{15}, 1/\sqrt{15})$$

Tenemos también $u + v + w = (7, 4, -1, 3)$ y $\|u + v + w\|^2 = 49 + 16 + 1 + 9 = 75$. Siendo así,

$$\|u\|^2 + \|v\|^2 + \|w\|^2 = 30 + 30 + 15 = 75 = \|u + v + w\|^2$$

con lo que se verifica el teorema de Pitágoras para el conjunto ortogonal S.

EJEMPLO 6.9. Consideremos el vector $u = (1, 1, 1, 1)$ en $\mathbf{R}^4$. Supongamos que queremos encontrar una base ortogonal de $u^\perp$, el complemento ortogonal de u. Observemos que $u^\perp$ es el espacio solución de la ecuación lineal.

$$x + y + z + t = 0 \qquad [1]$$

Hallamos una solución no nula v_1 de [1], digamos $v_1 = (0, 0, 1, -1)$. Queremos que nuestro segundo vector de la base v_2 sea solución de [1] y además ortogonal a v_1, esto es, que sea una solución del sistema

$$x + y + z + t = 0 \qquad z - t = 0 \qquad [2]$$

Hallamos una solución no nula v_2 de [2], digamos $v_2 = (0, 2, -1, 1)$. Queremos que nuestro tercer vector de la base v_3 sea solución de [1] y además ortogonal a v_1 y v_2, es decir, que sea una solución del sistema

$$x + y + z + t = 0 \qquad 2y - z - t = 0 \qquad z - t = 0 \qquad [3]$$

Hallamos una solución no nula de [3], por ejemplo $v_3 = (-3, 1, 1, 1)$. Entonces $\{v_1, v_2, v_3\}$ es una base ortogonal de $u^\perp$. (Obsérvese que hemos elegido las soluciones intermedias v_1 y v_2 de manera tal que cada nuevo sistema esté ya en forma escalonada. Esto hace los cálculos más sencillos.) Podemos determinar una base ortonormal de $u^\perp$ normalizando la base ortogonal de $u^\perp$ anterior. Tenemos

$$\|v_1\|^2 = 0 + 0 + 1 + 1 = 2 \qquad \|v_2\|^2 = 0 + 4 + 1 + 1 = 6 \qquad \|v_3\|^2 = 9 + 1 + 1 + 1 = 12$$

De modo que la siguiente es una base ortonormal de $u^\perp$.

$$v_1 = (0, 0, 1/\sqrt{2}, -1/\sqrt{2}) \qquad v_2 = (0, 2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6}) \qquad v_3 = (-3/\sqrt{12}, 1/\sqrt{12}, 1/\sqrt{12}, 1/\sqrt{12})$$

EJEMPLO 6.10. Supongamos que S consiste en los vectores de $\mathbf{R}^3$:

$$u_1 = (1, 2, 1) \qquad u_2 = (2, 1, -4) \qquad u_3 = (3, -2, 1)$$

En tal caso, S es ortogonal, pues u_1, u_2 y u_3 son ortogonales entre sí:

$$\langle u_1, u_2\rangle = 2 + 2 - 4 = 0 \qquad \langle u_1, u_3\rangle = 3 - 4 + 1 = 0 \qquad \langle u_2, u_3\rangle = 6 - 2 - 4 = 0$$

Así S es linealmente independiente y, dado que tiene tres elementos, es una base ortogonal de $\mathbf{R}^3$.

Supongamos ahora que queremos escribir $v = (4, 1, 18)$ como combinación lineal de u_1, u_2 y u_3. Comenzamos tomando v como combinación lineal de u_1, u_2 y u_3 usando incógnitas x, y, z como se hace a continuación:

$$(4, 1, 18) = x(1, 2, 1) + y(2, 1, -4) + z(3, -2, 1) \qquad [1]$$

Método 1. Desarrollamos [1] obteniendo

$$x + 2y + 3z = 4 \qquad 2x + y - 2z = 1 \qquad x - 4y + z = 18$$

de donde $x = 4$, $y = -3$, $z = 2$. De este modo, $v = 4u_1 - 3u_2 + 2u_3$.

Método 2. (Este método utiliza el hecho de que los vectores de la base son ortogonales, y la aritmética es mucho más simple.) Tomamos el producto interno de [1] y u_1 para conseguir

$$(4, 1, 18) \cdot (1, 2, 1) = x(1, 2, 1) \cdot (1, 2, 1) \quad \text{o} \quad 24 = 6x \quad \text{o} \quad x = 4$$

(Los dos últimos términos desaparecen porque u_1 es ortogonal a u_2 y a u_3.) Tomamos el producto interno de [1] y u_2 llegando a

$$(4, 1, 18) \cdot (2, 1, -4) = y(2, 1, -4) \cdot (2, 1, -4) \quad \text{o} \quad -63 = 21y \quad \text{o} \quad y = -3$$

Finalmente, tomamos el producto interno de [1] y u_3 para obtener

$$(4, 1, 18) \cdot (3, -2, 1) = z(3, -2, 1) \cdot (3, -2, 1) \quad \text{o} \quad 28 = 14z \quad \text{o} \quad z = 2$$

Así pues, $v = 4u_1 - 3u_2 + 2u_3$.

El procedimiento del Método 2 del Ejemplo 6.10 es válido en general; esto es,

Teorema 6.7: Supongamos que $\{u_1, u_2, \ldots, u_n\}$ es una base ortogonal de V. Para todo $v \in V$,

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle} u_n$$

(Véase el Problema 6.4 para su demostración.)

Nota: Cada escalar

$$k_i \equiv \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle} = \frac{\langle v, u_i\rangle}{\| u_i \|^2}$$

se denomina *coeficiente de Fourier* de v con respecto a los u_i, debido a que es análogo a un coeficiente en la serie de Fourier de una función. Este escalar tiene también una interpretación geométrica, que se discute más adelante.

PROYECCIONES

Consideremos un vector no nulo w en un espacio vectorial con producto interno V. Para todo $v \in V$ demostraremos (Problema 6.24) que

$$c = \frac{\langle v, w \rangle}{\langle w, w \rangle} = \frac{\langle v, w \rangle}{\| w \|^2}$$

es el único escalar tal que $v' = v - cw$ es ortogonal a w. La proyección de v a lo largo de w, como se indica en la Figura 6-5, se denota y define por

$$\text{proy}\,(v, w) = cw = \frac{\langle v, w \rangle}{\langle w, w \rangle}\, w = \frac{\langle v, w \rangle}{\| w \|^2}\, w$$

El escalar c también se llama el coeficiente de Fourier de v con respecto a w o la componente de v a lo largo de w.

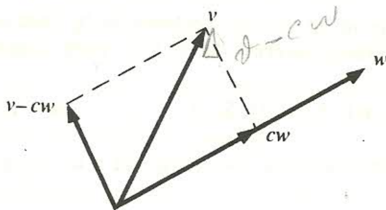


Figura 6-5.

EJEMPLO 6.11

a) Hallemos la compónente c y la proyección cw de $v = (1, 2, 3, 4)$ a lo largo de $w = (1, -3, 4, 2)$ en $\mathbf{R}^4$. Para ello comenzamos por calcular

$$\langle v, w \rangle = 1 - 6 + 12 - 8 = -1 \qquad \text{y} \qquad \| w \|^2 = 1 + 9 + 16 + 4 = 30$$

Entonces $c = -\frac{1}{30}$ y proy $(v, w) = cw = \left(-\frac{1}{30}, \frac{1}{10}, -\frac{2}{15}, \frac{1}{30}\right)$.

b) Sea V el espacio vectorial de los polinomios con producto interno $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. Hallemos la componente (coeficiente de Fourier) c y la proyección cg de $f(t) = 2t - 1$ a lo largo de $g(t) = t^2$. Empezamos calculando

$$\langle f, g \rangle = \int_0^1 (2t^3 - t^2)\, dt = \left[\frac{t^4}{2} - \frac{t^3}{3}\right]_0^1 = \frac{1}{6} \qquad \langle g, g \rangle = \int_0^1 t^4\, dt = \left[\frac{t^5}{5}\right]_0^1 = \frac{1}{5}$$

Entonces $c = \frac{5}{6}$ y proy $(f, g) = cg = 5t^2/6$.

La noción precedente puede generalizarse como sigue.

Teorema 6.8: Supongamos que $w_1, w_2, \ldots, w_r$ constituyen un conjunto ortogonal de vectores no nulos en V. Sea v un vector arbitrario en V. Definamos $v' = v - c_1 w_1 - c_2 w_2 - \cdots - c_r w_r$, siendo

$$c_1 = \frac{\langle v, w_1 \rangle}{\| w_1 \|^2}, \qquad c_2 = \frac{\langle v, w_2 \rangle}{\| w_2 \|^2}, \qquad \ldots, \qquad c_r = \frac{\langle v, w_r \rangle}{\| w_r \|^2}$$

En ese caso, v' es ortogonal a $w_1, w_2 \ldots, w_r$.

Nótese que los c_i en el teorema anterior son, respectivamente, las componentes (coeficientes de Fourier) de v a lo largo de los w_i. Además, el teorema enunciado a continuación (probado en el Problema 6.31) muestra que $c_1 w_1 + \cdots + c_r w_r$ es la aproximación más cercana a v por combinación lineal de $w_1, \ldots, w_r$.

Teorema 6.9: Supongamos que $w_1, w_2, \ldots, w_r$ forman un conjunto ortogonal de vectores no nulos en V. Sean v cualquier vector en V y c_i la componente de v a lo largo de w_i. Para escalares cualesquiera $a_1, \ldots, a_r$,

$$\left\| v - \sum_{k=1}^{r} c_k w_k \right\| \leq \left\| v - \sum_{k=1}^{r} a_k w_k \right\|$$

El siguiente teorema (demostrado en el Problema 6.32) se conoce como la *desigualdad de Bessel.*

Teorema 6.10: Supongamos que $\{e_1, e_2, \ldots, e_r\}$ es un conjunto ortonormal de vectores en V. Sean v cualquier vector de V y c_i el coeficiente de Fourier de v con respecto a e_i. Entonces

$$\sum_{k=1}^{r} c_k^2 \leq \| v \|^2$$

Nota: La noción de proyección incluye la de un vector a lo largo de un subespacio como se explica a continuación. Supongamos que W es un subespacio de V y que $v \in V$. Por el Teorema 6.4, $V = W \oplus W^{\perp}$; por consiguiente, v puede expresarse de forma única como

$$v = w + w' \qquad \text{con} \qquad w \in W, \quad w' \in W^{\perp}.$$

Llamamos a w la *proyección de v a lo largo de W*, lo que denotamos por $w = \text{proy}\,(v, W)$. (Véase la Figura 6-6). En particular, si $W = \text{lin}\,(w_1, \ldots, w_r)$, donde los w_i forman un conjunto ortogonal.

$$\text{proy}\,(v, W) = c_1 w_1 + c_2 w_2 + \cdots + c_r w_r$$

siendo c_i la componente de v a lo largo de w_i, como antes.

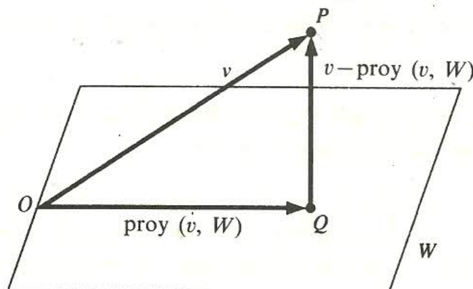


Figura 6-6.

6.6. PROCESO DE ORTOGONALIZACION DE GRAM-SCHMIDT

Supongamos que $\{v_1, v_2, \ldots, v_n\}$ es una base de un espacio con producto interno V. Podemos construir una base ortogonal $\{w_1, w_2, \ldots, w_n\}$ de V como sigue. Tomemos

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\| w_1 \|^2} w_1$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\| w_1 \|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\| w_2 \|^2} w_2$$

....

$$w_n = v_n - \frac{\langle v_n, w_1 \rangle}{\| w_1 \|^2} w_1 - \frac{\langle v_n, w_2 \rangle}{\| w_2 \|^2} w_2 - \cdots - \frac{\langle v_n, w_{n-1} \rangle}{\| w_{n-1} \|^2} w_{n-1}$$

Dicho de otro modo, para $k = 2, 3, \ldots, n$, definamos

$$w_k = v_k - c_{k1} w_1 - c_{k2} w_2 - \cdots - c_{k,k-1} w_{k-1}$$

donde $c_{ki} = \langle v_k, w_i \rangle / \|w_i\|^2$ es la componente de v_k a lo largo de w_i. De acuerdo con el Teorema 6.8, cada w_k es ortogonal a los precedentes, de modo que $w_1, w_2, \ldots, w_n$ forman una base ortogonal de V, como se pretendía. La normalización de cada w_k proporcionará una base ortonormal de V.

La construcción anterior se conoce como el *proceso de ortogonalización de Gram-Schmidt*. A continuación hacemos una serie de observaciones, por orden.

Nota 1: Cada vector w_k es combinación lineal de v_k y de los vectores w precedentes; por inducción, cada w_k es combinación lineal de $v_1, v_2, \ldots, v_k$.

Nota 2: Supongamos que $w_1, w_2, \ldots, w_r$ son linealmente independientes. Formarán, pues, una base de $U = \text{lin } w_i$. El proceso de ortogonalización de Gram-Schmidt sobre los w_i conduce a una base ortogonal de U.

Nota 3: En los cálculos manuales puede resultar más sencillo suprimir los denominadores en cada nuevo w_k multiplicando éste por un escalar apropiado, lo que no afecta a la ortogonalidad (Problema 6.75).

Los siguientes teoremas, demostrados en los Problemas 6.32 y 6.33, respectivamente, hacen uso del algoritmo y las notas que acabamos de ver.

Teorema 6.11: Sea $\{v_1, v_2, \ldots, v_n\}$ cualquier base de un espacio con producto interno V. Existe una base ortonormal $\{u_1, u_2, \ldots, u_n\}$ de V tal que la matriz de cambio de base desde $\{v_i\}$ hasta $\{u_i\}$ es triangular; es decir, para $k = 1, \ldots, n$,

$$u_k = a_{k1} v_1 + a_{k2} v_2 + \cdots + a_{kk} v_k$$

Teorema 6.12: Supongamos que $S = \{w_1, w_2, \ldots, w_r\}$ es una base ortogonal de un subespacio W de V. Entonces es posible extender S a una base ortogonal de V, o sea, pueden encontrarse vectores $w_{r+1}, \ldots, w_n$, tales que $\{w_1, w_2, \ldots, w_n\}$ sea una base de V.

EJEMPLO 6.12. Consideremos el subespacio U de $\mathbf{R}^4$ generado por

$$v_1 = (1, 1, 1, 1) \qquad v_2 = (1, 2, 4, 5) \qquad v_3 = (1, -3, -4, -2)$$

Hallemos una base ortonormal de U, encontrando primero una base ortogonal mediante el algoritmo de Gram-Schmidt. Primero tomamos $w_1 = v_1 = (1, 1, 1, 1)$. Después calculamos

$$v_2 - \frac{\langle v_2, w_1 \rangle}{\| w_1 \|^2} w_1 = (1, 2, 4, 5) - \frac{12}{4}(1, 1, 1, 1) = (-2, -1, 1, 2)$$

Tomamos $w_2 = (-2, -1, 1, 2)$. Luego hallamos

$$v_3 - \frac{\langle v_3, w_1 \rangle}{\| w_1 \|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\| w_2 \|^2} w_2 = (1, -3, -4, -2) - \frac{-8}{4}(1, 1, 1, 1) - \frac{-7}{10}(-2, -1, 1, 2) =$$
$$= \left(\frac{8}{5}, -\frac{17}{10}, -\frac{13}{10}, \frac{7}{5}\right)$$

Eliminamos los denominadores obteniendo $w_3 = (16, -17, -13, 14)$. Por último, normalizamos la base ortogonal

$$w_1 = (1, 1, 1, 1) \qquad w_2 = (-2, -1, 1, 2) \qquad w_3 = (16, -17, -13, 14)$$

Como $\|w_1\|^2 = 4$, $\|w_2\|^2 = 10$, $\|w_3\|^2 = 910$, una base ortonormal de U es:

$$u_1 = \frac{1}{2}(1, 1, 1, 1) \qquad u_2 = \frac{1}{\sqrt{10}}(-2, -1, 1, 2) \qquad u_3 = \frac{1}{\sqrt{910}}(16, -17, -13, 14)$$

EJEMPLO 6.13. Sea V el espacio vectorial de los polinomios $f(t)$ con producto interno $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)dt$. Aplicamos el algoritmo de Gram-Schmidt al conjunto $\{1, t, t^2, t^3\}$ para obtener una base ortogonal $\{f_0, f_1, f_2, f_3\}$, con coeficientes enteros, del subespacio U de los polinomios de grado $\leqslant 3$. Aquí utilizamos el hecho de que si $r + s = n$,

$$\langle t^r, t^s \rangle = \int_{-1}^{1} t^n \, dt = \left[\frac{t^{n+1}}{n+1}\right]_{-1}^{1} = \begin{cases} 2/(n+1) & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar} \end{cases}$$

Empezamos tomando $f_0 = 1$. Luego hallamos

$$t - \frac{\langle t, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 = t - \frac{0}{2} \cdot 1 = t$$

Sea $f_1 = t$. Entonces calculamos

$$t^2 - \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle t^2, t \rangle}{\langle t, t \rangle} t = t^2 - \frac{\frac{2}{3}}{2} \cdot 1 - \frac{0}{\frac{2}{3}} t = t^2 - \frac{1}{3}$$

Multiplicamos por 3 para llegar a $f_2 = 3t^2 - 1$. Después calculamos

$$t^3 - \frac{\langle t^3, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle t^3, t \rangle}{\langle t, t \rangle} t - \frac{\langle t^3, 3t^2 - 1 \rangle}{\langle 3t^2 - 1, 3t^2 - 1 \rangle}(3t^2 - 1) = t^3 - 0 \cdot 1 - \frac{\frac{2}{5}}{\frac{2}{3}} t - 0(3t^2 - 1) = t^3 - \frac{3}{5} t$$

Multiplicamos por 5 para llegar a $f_3 = 5t^3 - 3t$. Esto es, $\{1, t, 3t^2 - 1, 5t^3 - 3t\}$ es la base ortogonal de U requerida.

Nota: La normalización de los polinomios del Ejemplo 6.13 de forma que $p(1) = 1$ para todo polinomio $p(t)$ conduce a los

$$1,\ t,\ \tfrac{1}{2}(3t^2 - 1),\ \tfrac{1}{2}(5t^3 - 3t)$$

Estos son los cuatro primeros *polinomios de Legendre* (importantes en el estudio de ecuaciones diferenciales).

6.7. PRODUCTOS INTERNOS Y MATRICES

Esta sección investiga dos tipos de matrices que desempeñan un papel especial en la teoría de espacios reales con producto interno V: las matrices definidas positivas y las ortogonales. En este contexto, los vectores en $\mathbf{R}^n$ se representarán por vectores columna (de modo que $\langle u, v\rangle = u^T v$ denotará el producto interno usual en $\mathbf{R}^n$).

MATRICES DEFINIDAS POSITIVAS

Sea A una matriz real simétrica. Recordemos (Sección 4.11) que A es congruente a una matriz diagonal B, es decir, existe una matriz no singular P tal que $B = P^T AP$ es diagonal, así como que el número de entradas positivas de B es un invariante de A (Teorema 4.18, ley de inercia). Se dice que la matriz A es definida positiva si todas las entradas diagonales de B son positivas. Alternativamente se dice que A es definida positiva si $X^T AX > 0$ para todo vector no nulo X en $\mathbf{R}^n$.

EJEMPLO 6.14. Sea $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ -1 & -2 & 8 \end{pmatrix}$. Reducimos A a una matriz (congruente) diagonal (véase la Sección 4.11) efectuando $R_1 + R_3 \to R_3$ y la operación entre columnas correspondiente $C_1 + C_3 \to C_3$, y después $2R_2 + R_3 \to R_3$ y $2C_2 + C_3 \to C_3$:

$$A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Como la matriz diagonal tiene sólo entradas positivas, A es una matriz definida positiva.

Nota: Una matriz simétrica 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ es definida positiva si y sólo si las entradas diagonales a y d son positivas y el determinante $\det(A) = ad - bc = ad - b^2$ es positivo.

El teorema que sigue, demostrado en el Problema 6.40, es aplicable.

Teorema 6.13: Sea A una matriz real definida positiva. La función $\langle u, v\rangle = u^T Av$ es un producto interno en $\mathbf{R}^n$.

El Teorema 6.13 dice que toda matriz definida positiva A determina un producto interno. La discusión que ahora se inicia y el Teorema 6.15 pueden verse como el recíproco de este resultado.

Sean V un espacio con producto interno y $S = \{u_1, u_2, \ldots, u_n\}$ una base de V. La matriz A escrita a continuación se denomina *representación matricial* del producto interno en V relativa a la base S:

$$A = \begin{pmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle & \ldots & \langle u_1, u_n \rangle \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle & \ldots & \langle u_2, u_n \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle u_n, u_1 \rangle & \langle u_n, u_2 \rangle & \ldots & \langle u_n, u_n \rangle \end{pmatrix}$$

O sea, $A = (a_{ij})$, donde $a_{ij} = \langle u_i, u_j \rangle$.

Obsérvese que A es simétrica por serlo el producto interno, esto es, $\langle u_i, u_j \rangle = \langle u_j, u_i \rangle$. Asimismo, A depende tanto del producto interno en V como de la base S de V. Aún más, si S es una base ortogonal, A es necesariamente diagonal, y si S es una base ortonormal, A es la matriz identidad.

EJEMPLO 6.15. Los tres vectores siguientes forman una base S del espacio euclídeo $\mathbf{R}^3$:

$$u_1 = (1, 1, 0) \qquad u_2 = (1, 2, 3) \qquad u_3 = (1, 3, 5)$$

El cálculo de los $\langle u_i, u_j \rangle = \langle u_j, u_i \rangle$ proporciona:

$$\langle u_1, u_1 \rangle = 1 + 1 + 0 = 2 \qquad \langle u_1, u_2 \rangle = 1 + 2 + 0 = 3 \qquad \langle u_1, u_3 \rangle = 1 + 3 + 0 = 4$$

$$\langle u_2, u_2 \rangle = 1 + 4 + 9 = 14 \qquad \langle u_2, u_3 \rangle = 1 + 6 + 15 = 22 \qquad \langle u_3, u_3 \rangle = 1 + 9 + 25 = 35$$

De este modo,

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 14 & 22 \\ 4 & 22 & 35 \end{pmatrix}$$

es la representación matricial del producto interno usual en $\mathbf{R}^3$ relativa a la base S.

Son aplicables los teoremas enunciados a continuación y demostrados en los Problemas 6.41 y 6.42, respectivamente.

Teorema 6.14: Sea A la representación matricial de un producto interno relativa a una base S de V. Para todo par de vectores $u, v \in V$ tenemos

$$\langle u, v \rangle = [u]^T A [v]$$

donde $[u]$ y $[v]$ denotan los vectores (columna) coordenados respecto a la base S.

Teorema 6.15: Sea A la representación matricial de cualquier producto interno en V. Entonces A es una matriz definida positiva.

MATRICES ORTOGONALES

Recordemos (Sección 4.6) que una matriz P es ortogonal si es no singular y $P^{-1} = P^T$, esto es, si $PP^T = P^T P = I$. Esta subsección lleva más lejos la investigación de dichas matrices. Empezamos recordando (Teorema 4.5) una importante caracterización de las matrices ortogonales.

Teorema 6.16: Sea P una matriz real. Son equivalentes las tres propiedades:

i) P es ortogonal, es decir, $P^T = P^{-1}$.

ii) Las filas de P forman un conjunto ortonormal de vectores.

iii) Las columnas de P forman un conjunto ortonormal de vectores.

(El teorema precedente sólo es cierto cuando se utiliza el producto interno usual de $\mathbf{R}^n$, no cuando $\mathbf{R}^n$ está dotado de cualquier otro producto interno.)

Nota: Toda matriz ortogonal 2×2 presenta las formas $\begin{pmatrix} \cos\theta & \text{sen}\,\theta \\ -\text{sen}\,\theta & \cos\theta \end{pmatrix}$ o $\begin{pmatrix} \cos\theta & \text{sen}\,\theta \\ \text{sen}\,\theta & -\cos\theta \end{pmatrix}$ para algún número real θ (Teorema 4.6).

EJEMPLO 6.16

Sea $P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{pmatrix}$. Las filas son ortogonales entre sí y son vectores unitarios, o sea, constituyen un conjunto ortonormal de vectores. Podemos decir, pues, que P es ortogonal.

Los dos teoremas siguientes, demostrados en los Problemas 6.48 y 6.49, respectivamente, muestran algunas relaciones importantes entre matrices ortogonales y bases ortonormales de un espacio con producto interno V.

Teorema 6.17: Supongamos que $E = \{e_i\}$ y $E' = \{e'_i\}$ son bases ortonormales en un espacio con producto interno V. Sea P la matriz de cambio de base desde E hasta E'. En ese caso, P es ortogonal.

Teorema 6.18: Sean $\{e_1, \ldots, e_n\}$ una base ortonormal de un espacio con producto interno V y $P = (a_{ij})$ una matriz ortogonal. Entonces los n vectores que siguen forman una base ortonormal de V:

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n \qquad (i = 1, 2, \ldots, n)$$

6.8. ESPACIOS COMPLEJOS CON PRODUCTO INTERNO

Esta sección considera espacios vectoriales sobre el cuerpo complejo $\mathbf{C}$. Recordemos primero algunas de las propiedades de los números complejos (Sección 2.9). Supongamos que $z \in \mathbf{C}$, es decir, $z = a + bi$, donde $a, b \in \mathbf{R}$. En tal caso,

$$\bar{z} = a - bi \qquad z\bar{z} = a^2 + b^2 \qquad \text{y} \qquad |z| = \sqrt{a^2 + b^2}$$

Asimismo, para $z, z_1, z_2 \in \mathbf{C}$,

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \qquad \overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2} \qquad \bar{\bar{z}} = z$$

y z es real si y sólo si $\bar{z} = z$.

Definición: Sea V un espacio vectorial sobre **C**. Supongamos que a cada par de vectores $u, v \in V$ se le asigna un número complejo, denotado por $\langle u, v \rangle$. Esta función se denomina un *producto interno* (*complejo*) en V si satisface los axiomas:

$[I_1^*]$ (Propiedad lineal) $\langle au_1 + bu_2, v \rangle = a\langle u_1, v \rangle + b\langle u_2, v \rangle$.

$[I_2^*]$ (Propiedad simétrica conjugada) $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

$[I_3^*]$ (Propiedad definida positiva) $\langle u, u \rangle \geqslant 0$; y $\langle u, u \rangle = 0$ si y sólo si $u = 0$.

El espacio V sobre **C** recibe entonces el nombre de *espacio* (*complejo*) *con producto interno.*

Observemos que un producto interno complejo difiere sólo ligeramente de un producto interno real (únicamente $[I_2^*]$ difiere de $[I_2]$). De hecho, muchas de las definiciones y propiedades de un espacio complejo con producto interno coinciden con las de un espacio con producto interno real. Sin embargo, algunas de las demostraciones deben adaptarse al caso complejo.

El axioma $[I_1^*]$ es también equivalente a las dos condiciones:

$$a) \quad \langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle \qquad \text{y} \qquad b) \quad \langle ku, v \rangle = k\langle u, v \rangle$$

En cambio,

$$\langle u, kv \rangle = \overline{\langle kv, u \rangle} = \overline{k\langle v, u \rangle} = \bar{k}\overline{\langle v, u \rangle} = \bar{k}\langle u, v \rangle$$

(En otras palabras, debemos tomar el conjugado de un escalar complejo cuando éste se saca de la segunda posición del producto interno.) Demostraremos (Problema 6.50) que, de hecho, el producto interno es *antilineal* en la segunda posición, esto es,

$$\langle u, av_1 + bv_2 \rangle = \bar{a}\langle u, v_1 \rangle + \bar{b}\langle u, v_2 \rangle$$

Puede probarse (Problema 6.95), análogamente, que

$$\langle a_1u_1 + a_2u_2, b_1v_1 + b_2v_2 \rangle = a_1\bar{b}_1\langle u_1, v_1 \rangle + a_1\bar{b}_2\langle u_1, v_2 \rangle + a_2\bar{b}_1\langle u_2, v_1 \rangle + a_2\bar{b}_2\langle u_2, v_2 \rangle$$

y, por inducción,

$$\left\langle \sum_{i=1}^{m} a_i u_i, \sum_{j=1}^{n} b_j v_j \right\rangle = \sum_{i,j} a_i \bar{b}_j \langle u_i, v_j \rangle$$

Podemos hacer, por orden, las observaciones similares:

Nota 1: El axioma $[I_1^*]$ por sí mismo implica $\langle 0, 0 \rangle = \langle 0v, 0 \rangle = 0\langle v, 0 \rangle = 0$. En consecuencia, $[I_1^*]$, $[I_2^*]$ e $[I_3^*]$ son equivalentes a $[I_1^*]$, $[I_2^*]$ y el axioma:

$$[I_3^{*\prime}] \quad \text{Si } u \neq 0, \quad \text{necesariamente} \quad \langle u, u \rangle > 0$$

O sea, una función que satisface $[I_1^*]$, $[I_2^*]$ e $[I_3^{*\prime}]$ es un producto interno (complejo) en V.

Nota 2: De acuerdo con $[I_2^*]$, $\langle u, u \rangle = \overline{\langle u, u \rangle}$. Así $\langle u, u \rangle$ debe ser real. Por $[I_3^*]$, $\langle u, u \rangle$ debe ser no negativo y por tanto existirá su raíz cuadrada. Como hacíamos con los espacios con producto interno reales, definimos la norma o longitud de u como $\|u\| = \sqrt{\langle u, u \rangle}$.

Nota 3: Junto a la norma, definimos las nociones de ortogonalidad, complemento ortogonal, y conjuntos ortogonales y ortonormales como antes. De hecho, las definiciones de distancia, coeficiente de Fourier y proyección son las mismas que en el caso real.

EJEMPLO 6.17. Sean $u = (z_i)$ y $v = (w_i)$ vectores en $\mathbf{C}^n$. Entonces

$$\langle u, v \rangle = \sum_{k=1}^{n} z_k \bar{w}_k = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n$$

es un producto interno en $\mathbf{C}^n$ llamado el producto usual o convencional en $\mathbf{C}^n$. (Supondremos este producto interno en $\mathbf{C}^n$ a menos que se establezca o sobrentienda otra cosa.) En caso de que u y v sean reales, tenemos $\bar{w}_i = w_i$ y

$$\langle u, v \rangle = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n = z_1 w_1 + z_2 w_2 + \cdots + z_n w_n$$

Dicho de otro modo, este producto interno se reduce a su análogo en $\mathbf{R}^n$ cuando las entradas son reales.

Nota: Aceptando que u y v son vectores columna, el producto interno precedente puede definirse según $\langle u, v \rangle = u^T \bar{v}$.

EJEMPLO 6.18

a) Sea V el espacio vectorial de las funciones complejas continuas definidas sobre el intervalo (real) $a \leqslant t \leqslant b$. El producto interno usual en V es:

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}\, dt$$

b) Sea U el espacio vectorial de las matrices $m \times n$ sobre $\mathbf{C}$. Supongamos que $A = (z_{ij})$ y $B = (w_{ij})$ son elementos de U. El producto interno usual en U es:

$$\langle A, B \rangle = \operatorname{tr} B^H A = \sum_{i=1}^{m} \sum_{j=1}^{n} \bar{w}_{ij} z_{ij}$$

Como de costumbre, $B^H = \bar{B}^T$, o sea, B^H es la traspuesta conjugada de B.

A continuación se presenta una lista de teoremas para espacios complejos con producto interno, análogos a los dados para el caso real (el Teorema 6.19 se demuestra en el Problema 6.53).

Teorema 6.19 (Cauchy-Schwarz): Sea V un espacio complejo con producto interno. Entonces

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|$$

Teorema 6.20: Sea W un subespacio de un espacio complejo con producto interno V. En tal caso, $V = W \oplus W$.⊥

Teorema 6.21: Supongamos que $\{u_1, u_2, \ldots, u_n\}$ es una base ortogonal de un espacio vectorial complejo V. Para todo $v \in V$,

$$v = \frac{\langle v, u_1 \rangle}{\|u_1\|^2} u_1 + \frac{\langle v, u_2 \rangle}{\|u_2\|^2} u_2 + \cdots + \frac{\langle v, u_n \rangle}{\|u_n\|^2} u_n$$

Teorema 6.22: Supongamos que $\{u_1, u_2, \ldots, u_n\}$ es una base de un espacio complejo con producto interno V. Sea $A = (a_{ij})$ la matriz compleja definida por $a_{ij} = \langle u_i, u_j \rangle$. Para todo par de vectores $u, v \in V$,

$$\langle u, v \rangle = [u]^T A \overline{[v]}$$

donde $[u]$ y $[v]$ son los vectores columna coordenados en la base $\{u_i\}$ dada. (*Nota:* Se dice que esta matriz A representa el producto interno en V.)

Teorema 6.23: Sea A una matriz hermítica (esto es, $A^H = \bar{A}^T = A$) tal que $X^T A \bar{X}$ es real y positivo para todo vector no nulo $X \in \mathbf{C}^n$. En tal caso, $\langle u, v \rangle = u^T A \bar{v}$ es un producto interno en $\mathbf{C}^n$.

Teorema 6.24: Sea A la matriz que representa un producto interno en V. Entonces A es hermítica y $X^T A \bar{X}$ es real y positivo para cualquier vector no nulo en $\mathbf{C}^n$.

6.9. ESPACIOS VECTORIALES NORMADOS

Empecemos con una definición.

Definición: Sea V un espacio vectorial real o complejo. Supongamos que a cada $v \in V$ se le asigna un número real, denotado por $\|v\|$. Esta función $\|\cdot\|$ se llama una *norma* en V si satisface los axiomas:

$[N_1]$ $\|v\| \geqslant 0$; y $\|v\| = 0$ si y sólo si $v = 0$.

$[N_2]$ $\|kv\| = |k| \, \|v\|$.

$[N_3]$ $\|u + v\| \leqslant \|u\| + \|v\|$.

El espacio vectorial V con una norma se denomina *espacio vectorial normado.*

Se pueden hacer, por orden, una serie de observaciones.

Nota 1: El axioma $[N_2]$ por sí mismo implica $\|0\| = \|0v\| = 0\|v\| = 0$. De acuerdo con esto, $[N_1]$, $[N_2]$ y $[N_3]$ son equivalentes a $[N_2]$, $[N_3]$ y el axioma:

$$[N'_1] \quad \text{Si } v \neq 0, \quad \text{necesariamente} \quad \|v\| > 0$$

Es decir, una función $\|\cdot\|$ que satisfaga $[N'_1$, $[N_2]$ y $[N_3]$ es una norma en un espacio vectorial V.

Nota 2: Supongamos que V es un espacio con producto interno. La norma en V definida por $\|v\| = \sqrt{\langle v, v \rangle}$ satisface $[N_1]$, $[N_2]$ y $[N_3]$. Así todo espacio con producto interno V es un espacio vectorial normado. No obstante, puede haber normas en un espacio vectorial V que no provengan de un producto interno.

Nota 3: Sea V un espacio vectorial normado. La *distancia* entre los vectores $u, v \in V$ se denota y define por $d(u, v) = \|u - v\|$.

El teorema enunciado a continuación revela la razón principal por la que $d(u, v)$ recibe el nombre de distancia entre u y v.

Teorema 6.25: Sea V un espacio vectorial normado. La función $d(u, v) = \|u - v\|$ satisface los tres axiomas de un espacio métrico:

$[M_1]$ $d(u, v) \geqslant 0$; y $d(u, v) = 0$ si y sólo si $u = v$.

$[M_2]$ $d(u, v) = d(v, u)$.

$[M_3]$ $d(u, v) \leqslant d(u, w) + d(w, v)$.

NORMAS EN $\mathbf{R}^n$ Y $\mathbf{C}^n$

Las siguientes definiciones corresponden a tres importantes normas en $\mathbf{R}^n$ y $\mathbf{C}^n$:

$$\|(a_1, \ldots, a_n)\|_\infty = \text{máx}\ (|a_i|)$$
$$\|(a_1, \ldots, a_n)\|_1 = |a_1| + |a_2| + \cdots + |a_n|$$
$$\|(a_1, \ldots, a_n)\|_2 = \sqrt{|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2}$$

(Nótese que se utilizan subíndices para distinguir las tres normas.) Las normas $\|\cdot\|_\infty$, $\|\cdot\|_1$ y $\|\cdot\|_2$ se denominan *norma uniforme* (o *del supremo*), 1-*norma* y 2-*norma*, respectivamente. Observemos que $\|\cdot\|_2$ es la norma en $\mathbf{R}^n(\mathbf{C}^n)$ inducida por el producto interno usual. (Denotaremos por d_∞, d_1 y d_2 las correspondientes funciones distancia.)

EJEMPLO 6.19. Consideremos los vectores $u = (1, -5, 3)$ y $v = (4, 2, -3)$ en $\mathbf{R}^3$.

a) La norma uniforme elige el máximo de los valores absolutos de las componentes, luego

$$\|u\|_\infty = 5 \quad \text{y} \quad \|v\|_\infty = 4$$

b) La 1-norma suma los valores absolutos de las componentes. De este modo,

$$\|u\|_1 = 1 + 5 + 3 = 9 \quad \text{y} \quad \|v\|_1 = 4 + 2 + 3 = 9$$

c) La 2-norma es igual a la raíz cuadrada de la suma de los cuadrados de las componentes (o sea, la norma inducida por el producto interno usual en $\mathbf{R}^3$). Así pues,

$$\|u\|_2 = \sqrt{1 + 25 + 9} = \sqrt{35} \quad \text{y} \quad \|v\|_2 = \sqrt{16 + 4 + 9} = \sqrt{29}$$

d) Dado que $u - v = (1 - 4, -5 - 2, 3 + 3) = (-3, -7, 6)$, tenemos

$$d_\infty(u, v) = 7 \qquad d_1(u, v) = 3 + 7 + 6 = 16 \qquad d_2(u, v) = \sqrt{9 + 49 + 36} = \sqrt{94}$$

EJEMPLO 6.20. Consideremos el plano cartesiano $\mathbf{R}^2$ mostrado en la Figura 6-7.

a) Sea D_1 el conjunto de puntos $u = (x, y)$ en $\mathbf{R}^2$ tales que $\|u\|_2 = 1$. D_1 consiste, pues, en los puntos (x, y) tales que $\|u\|_2^2 = x^2 + y^2 = 1$. Siendo así, D_1 es el círculo unidad, como se muestra en la Figura 6-7.

b) Sea D_2 el conjunto de puntos $u = (x, y)$ en $\mathbf{R}^2$ tales que $\|u\|_1 = 1$. Entonces D_2 consiste en los puntos (x, y) tales que $\|u\|_1 = |x| + |y| = 1$. De este modo, D_2 es el rombo inscrito en el círculo unidad, como se ilustra en la Figura 6-7.

c) Sea D_3 el conjunto de puntos $u = (x, y)$ en $\mathbf{R}^2$ tales que $\|u\|_\infty = 1$. En ese caso, D_3 consiste en los puntos (x, y) tales que $\|u\|_\infty = \text{máx}\,(|x|, |y|) = 1$. Así pues, D_3 es el cuadrado circunscrito al círculo unidad, como se indica en la Figura 6-7.

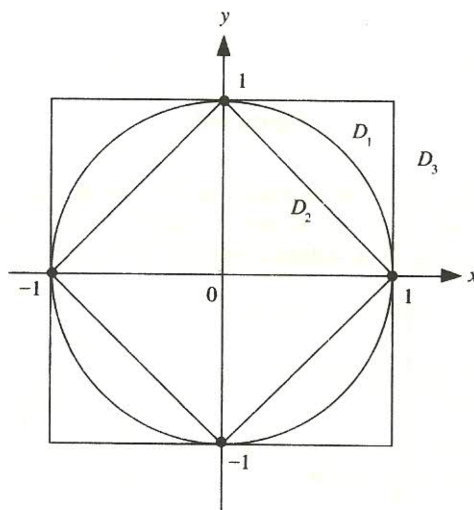


Figura 6-7.

NORMAS EN $C[a, b]$

Consideremos el espacio vectorial $V = C[a, b]$ de las funciones continuas en el intervalo $a \leqslant t \leqslant b$. Recordemos que un producto interno en V es:

$$\langle f, g \rangle = \int_a^b f(t)g(t)\,dt$$

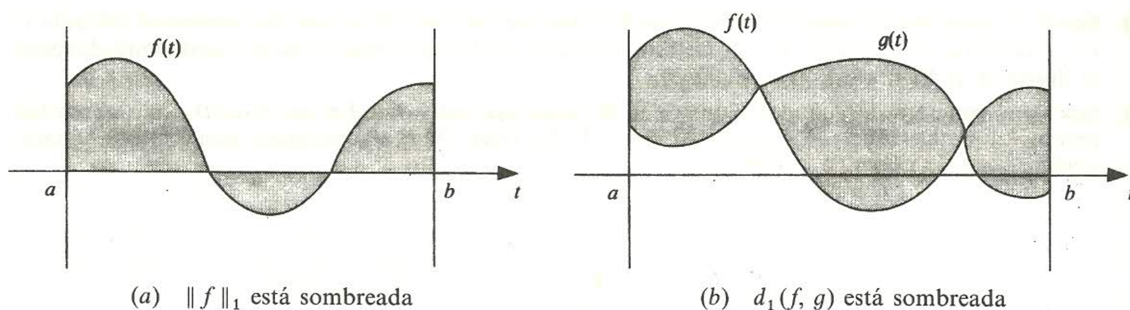
El producto interno anterior define la norma en $V = C[a, b]$ (análoga a la $\|\cdot\|_2$ en $\mathbf{R}^n$):

$$\|f\|_2 = \int_a^b [f(t)]^2\,dt$$

El ejemplo que sigue define otras dos normas en $V = C[a, b]$.

EJEMPLO 6.21

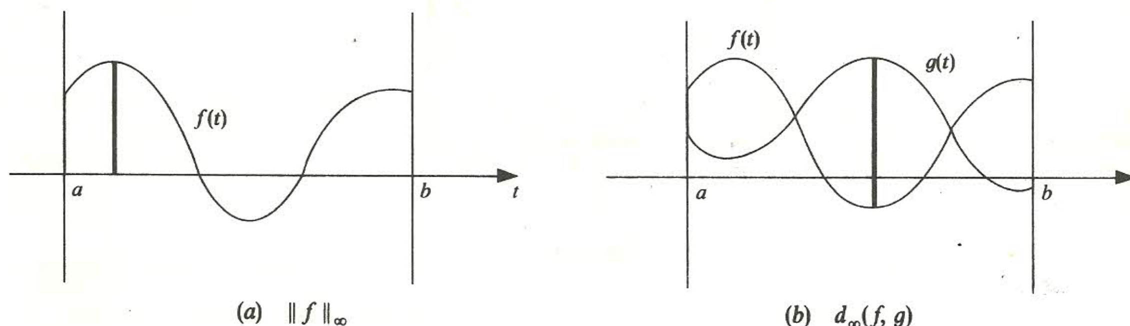
a) Sea $\|f\|_1 = \int_a^b |f(t)|\,dt$. (Esta norma es análoga a la $\|\cdot\|_1$ en $\mathbf{R}^n$.) Existe una descripción geométrica de $\|f\|_1$ y la distancia $d_1(f, g)$. Como se ilustra en la Figura 6-8, $\|f\|_1$ es el área entre la función $|f|$ y el eje t, mientras que $d_1(f, g)$ es el área entre las funciones f y g.



(*a*) $\|f\|_1$ está sombreada (*b*) $d_1(f, g)$ está sombreada

Figura 6-8.

b) Sea $\|f\|_\infty = \text{máx}\,(|f(t)|)$. (Esta norma es análoga a la $\|\cdot\|_\infty$ en $\mathbf{R}^n$.) Existe una descripción geométrica de $\|f\|_\infty$ y la función distancia $d_\infty(f, g)$. Como se muestra en la Figura 6-9, $\|f\|_\infty$ es la distancia máxima entre f y el eje t y $d_\infty(f, g)$ la distancia máxima entre f y g.



(*a*) $\|f\|_\infty$ (*b*) $d_\infty(f, g)$

Figura 6-9.

PROBLEMAS RESUELTOS

PRODUCTOS INTERNOS

6.1. Desarrollar $\langle 5u_1 + 8u_2, 6v_1 - 7v_2\rangle$.

Usando la linealidad en las dos posiciones,

$$\begin{aligned}\langle 5u_1 + 8u_2, 6v_1 - 7v_2\rangle &= \langle 5u_1, 6v_1\rangle + \langle 5u_1, -7v_2\rangle + \langle 8u_2, 6v_1\rangle + \langle 8u_2, -7v_2\rangle = \\ &= 30\langle u_1, v_1\rangle - 35\langle u_1, v_2\rangle + 48\langle u_2, v_1\rangle - 56\langle u_2, v_2\rangle\end{aligned}$$

[*Nota:* Obsérvese la similitud entre el desarrollo anterior y el de $(5a + 8b)(6c - 7d)$ en el álgebra ordinaria.]

6.2. Considérense los vectores en $\mathbf{R}^3$: $u = (1, 2, 4)$, $v = (2, -3, 5)$, $w = (4, 2, -3)$. Hallar: *a*) $u \cdot v$, *b*) $u \cdot w$, *c*) $v \cdot w$, *d*) $(u + v) \cdot W$, *e*) $\|u\|$, *f*) $\|v\|$, *g*) $\|u + v\|$.

a) Multiplicamos las componentes correspondientes y las sumamos obteniendo $u \cdot v = 2 - 6 + 20 = 16$.

b) $u \cdot w = 4 + 4 - 12 = -4$.

c) $v \cdot w = 8 - 6 - 15 = -13$.

d) Calculamos primero $u + v = (3, -1, 9)$. En ese caso, $(u + v) \cdot w = 12 - 2 - 27 = -17$. Alternativamente, usando $[I_1]$, $(u + v) \cdot w = u \cdot w + v \cdot w = -4 - 13 = -17$.

e) Calculamos primero $\|u\|^2$ elevando al cuadrado las componentes de u y sumando:

$$\|u\|^2 = 1^2 + 2^2 + 4^2 = 1 + 4 + 16 = 21 \qquad \text{y así} \qquad \|u\| = \sqrt{21}$$

f) $\|v\|^2 = 4 + 9 + 25 = 38$ y así $\|v\| = \sqrt{38}$.

g) Según *d*), $u + v = (3, -1, 9)$. De aquí $\|u + v\|^2 = 9 + 1 + 81 = 91$ y $\|u + v\| = \sqrt{91}$.

6.3. Verificar que el siguiente es un producto interno en $\mathbf{R}^2$:

$$\langle u, v\rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2 \qquad \text{donde} \qquad u = (x_1, x_2),\ v = (y_1, y_2)$$

Método 1. Comprobamos que se satisfacen los tres axiomas de un producto interno. Tomando $w = (z_1, z_2)$ encontramos

$$au + bw = a(x_1, x_2) + b(z_1, z_2) = (ax_1 + bz_1, ax_2 + bz_2)$$

Así pues,

$$\begin{aligned}\langle au + bw, v\rangle &= \langle (ax_1, ax_2 + bz_2), (y_1, y_2)\rangle = \\ &= (ax_1 + bz_1)y_1 - (ax_1 + bz_1)y_2 - (ax_2 + bz_2)y_1 + 3(ax_2 + bz_2)y_2 = \\ &= a(x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2) + b(z_1 y_1 - z_1 y_2 - z_2 y_1 + 3z_2 y_2) = \\ &= a\langle u, v\rangle + b\langle w, v\rangle\end{aligned}$$

y por tanto se satisface el axioma $[I_1]$. Asimismo,

$$\langle v, u\rangle = y_1 x_1 - y_1 x_2 - y_2 x_1 + 3y_2 x_2 = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2 = \langle u, v\rangle$$

y se satisface $[I_2]$. Finalmente,

$$\langle u, u\rangle = x_1^2 - 2x_1 x_2 + 3x_2^2 = x_1^2 - 2x_1 x_2 + x_2^2 + 2x_2^2 = (x_1 - x_2)^2 + 2x_2^2 \geq 0$$

Además, $\langle u, u\rangle = 0$ si y sólo si $x_1 = 0$, $x_2 = 0$, es decir, $u = 0$. Por consiguiente, se satisface el último axioma $[I_3]$.

Método 2. Razonamos con matrices. Esto es, escribimos $\langle u, v\rangle$ en notación matricial:

$$\langle u, v\rangle = u^T A v = (x_1, x_2)\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Como A es real y simétrica, sólo necesitamos probar que es definida positiva. Efectuando la operación elemental entre filas $R_1 + R_2 \to R_2$, seguida de la correspondiente operación elemental entre columnas $C_1 + C_2 \to C_2$, llevamos A a la forma diagonal $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. De este modo, A es definida positiva, de acuerdo con lo cual, $\langle u, v\rangle$ es un producto interno.

6.4. Considérense los vectores $u = (1, 5)$ y $v = (3, 4)$ en $\mathbf{R}^2$. Hallar:

a) $\langle u, v\rangle$ con respecto al producto interno usual en $\mathbf{R}^2$.

b) $\langle u, v\rangle$ con respecto al producto interno en $\mathbf{R}^2$ del Problema 6.3.

c) $\|v\|$ empleando el producto interno usual en $\mathbf{R}^2$.

d) $\|v\|$ empleando el producto interno en $\mathbf{R}^2$ del Problema 6.3.

a) $\langle u, v\rangle = 3 + 20 = 23$.

b) $\langle u, v\rangle = 1 \cdot 3 - 1 \cdot 4 - 5 \cdot 3 + 3 \cdot 5 \cdot 4 = 3 - 4 - 15 + 60 = 44$.

c) $\|v\|^2 = \langle v, v\rangle = \langle (3, 4), (3, 4)\rangle = 9 + 16 = 25$, luego $\|v\| = 5$.

d) $\|v\|^2 = \langle v, v\rangle = \langle (3, 4), (3, 4)\rangle = 9 - 12 - 12 + 48 = 33$; luego $\|v\| = \sqrt{33}$.

6.5. Considérense el espacio vectorial V de los polinomios con producto interno definido por $\int_0^1 f(t)g(t)dt$ y los polinomios $f(t) = t + 2$, $g(t) = 3t - 2$ y $h(t) = t^2 - 2t - 3$. Hallar: *a)* $\langle f, g\rangle$ y $\langle f, h\rangle$, *b)* $\|f\|$ y $\|g\|$. *c)* Normalizar f y g.

a) Integramos como sigue:

$$\langle f, g\rangle = \int_0^1 (t + 2)(3t - 2)\, dt = \int_0^1 (3t^2 + 4t - 4)\, dt = [t^3 + 2t^2 - 4t]_0^1 = -1$$

$$\langle f, h\rangle = \int_0^1 (t + 2)(t^2 - 2t - 3)\, dt = \left[\frac{t^4}{4} - \frac{7t^2}{2} - 6t\right]_0^1 = -\frac{37}{4}$$

b)

$$\langle f, f\rangle = \int_0^1 (t + 2)(t + 2)\, dt = \frac{19}{3} \quad \text{y} \quad \|f\| = \sqrt{\langle f, f\rangle} + \sqrt{19/3} = \frac{\sqrt{57}}{3}\ .$$

$$\langle g, g\rangle = \int_0^1 (3t - 2)(3t - 2) = 1;\ \text{por tanto,}\ \|g\| = \sqrt{1} = 1$$

c) Siendo $\|f\| = \dfrac{\sqrt{57}}{3}$, $\hat{f} = \dfrac{1}{\|f\|} f = \dfrac{3}{\sqrt{57}}(t + 2)$. Nótese que g ya es un vector unitario, pues $\|g\| = 1$; por consiguiente, $\hat{g} = g = 3t - 2$.

6.6. Sea V el espacio vectorial de las matrices reales 2×3 con el producto interno $\langle A, B\rangle = \operatorname{tr} B^T A$ y considérense las matrices

$$A = \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad C = \begin{pmatrix} 3 & -5 & 2 \\ 1 & 0 & -4 \end{pmatrix}$$

Calcular: *a)* $\langle A, B\rangle$, $\langle A, C\rangle$ y $\langle B, C\rangle$; *b)* $\langle 2A + 3B, 4C\rangle$; *c)* $\|A\|$ y $\|B\|$. *d)* Normalizar A y B.

a) [Usamos $\langle A, B\rangle = \operatorname{tr} B^T A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}$, la suma de los productos de las entradas correspondientes.]

$$\langle A, B\rangle = 9 + 16 + 21 + 24 + 25 + 24 = 119$$
$$\langle A, C\rangle = 27 - 40 + 14 + 6 + 0 - 16 = -9$$
$$\langle B, C\rangle = 3 - 10 + 6 + 4 + 0 - 24 = -21$$

b) Hallamos $2A + 3B = \begin{pmatrix} 21 & 22 & 23 \\ 24 & 25 & 26 \end{pmatrix}$ y $4C = \begin{pmatrix} 12 & -20 & 8 \\ 4 & 0 & -16 \end{pmatrix}$. Entonces

$$\langle 2A + 3B, 4C\rangle = 252 - 440 + 184 + 96 + 0 - 416 = -324$$

Alternativamente, usando la propiedad lineal de los productos internos,

$$\langle 2A + 3B, 4C\rangle = 8\langle A, C\rangle + 12\langle B, C\rangle = 8(-9) + 12(-21) = -324$$

c) $\left[\text{Utilizamos } \|A\|^2 = \langle A, A\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2, \text{ la suma de los cuadrados de los elementos de } A.\right]$

$$\|A\|^2 = \langle A, A\rangle = 9^2 + 8^2 + 7^2 + 6^2 + 5^2 + 4^2 = 271 \quad \text{y así} \quad \|A\| = \sqrt{271}$$
$$\|B\|^2 = \langle B, B\rangle = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 = 91 \quad \text{y así} \quad \|B\| = \sqrt{91}$$

d)
$$\hat{A} = \frac{1}{\|A\|} A = \frac{1}{\sqrt{271}} A = \begin{pmatrix} 9/\sqrt{271} & 8/\sqrt{271} & 7/\sqrt{271} \\ 6/\sqrt{271} & 5/\sqrt{271} & 4/\sqrt{271} \end{pmatrix}$$
$$\hat{B} = \frac{1}{\|B\|} B = \frac{1}{\sqrt{91}} B = \begin{pmatrix} 1/\sqrt{91} & 2/\sqrt{91} & 3/\sqrt{91} \\ 4/\sqrt{91} & 5/\sqrt{91} & 6/\sqrt{91} \end{pmatrix}$$

6.7. Determinar la distancia $d(u, v)$ entre los vectores:

a) $u = (1, 3, 5, 7)$ y $v = (4, -2, 8, 1)$ en $\mathbf{R}^4$.

b) $u = t + 2$ y $v = 3t - 2$, donde $\langle u, v\rangle = \int_0^1 u(t)v(t)dt$.

Usamos $d(u, v) = \|u - v\|$.

a) $u - v = (-3, 5, -3, 6)$. De este modo,

$$\|u - v\|^2 = 9 + 25 + 9 + 36 = 79 \quad \text{por lo que} \quad d(u, v) = \|u - v\| = \sqrt{79}$$

b) $u - v = -2t + 4$. De aquí

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \int_0^1 (-2t + 4)(-2t + 4)\, dt =$$
$$= \int_0^1 (4t^2 - 16t + 16)\, dt = \left[\frac{4}{3}t^3 - 8t^2 + 16t\right]_0^1 = \frac{28}{3}$$

Así pues, $d(u, v) = \sqrt{\frac{28}{3}} = \frac{2}{3}\sqrt{21}$.

6.8. Hallar $\cos\theta$, siendo θ el ángulo entre:

a) $u = (1, -3, 2)$ y $v = (2, 1, 5)$ en $\mathbf{R}^3$.

b) $u = (1, 3, -5, 4)$ y $v = (2, -3, 4, 1)$ en $\mathbf{R}^4$.

c) $f(t) = 2t - 1$ y $g(t) = t^2$, donde $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$.

d) $A = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix}$, donde $\langle A, B\rangle = \text{tr } B^T A$.

Usamos $\cos\theta = \dfrac{\langle u, v\rangle}{\|u\|\,\|v\|}$.

a) Calculamos $\langle u, v\rangle = 2 - 3 + 10 = 9$, $\|u\|^2 = 1 + 9 + 4 = 14$, $\|v\|^2 = 4 + 1 + 25 = 30$. Siendo así,

$$\cos\theta = \frac{9}{\sqrt{14}\sqrt{30}} = \frac{9}{2\sqrt{105}}$$

b) Aquí $\langle u, v\rangle = 2 - 9 - 20 + 4 = -23$, $\|u\|^2 = 1 + 9 + 25 + 16 = 51$, $\|v\|^2 = 4 + 9 + 16 + 1 = 30$. De este modo,

$$\cos\theta = \frac{-23}{\sqrt{51}\sqrt{30}} = \frac{-23}{3\sqrt{170}}$$

c) Calculamos

$$\langle f, g\rangle = \int_0^1 (2t^3 - t^2)\,dt = \left[\frac{t^4}{2} - \frac{t^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 (4t^2 - 4t + 1)\,dt = \frac{1}{3} \qquad \text{y} \qquad \|g\|^2 = \langle g, g\rangle = \int_0^1 t^4\,dt = \frac{1}{3}$$

Por tanto, $\theta = \dfrac{\frac{1}{6}}{(1/\sqrt{3})(1/\sqrt{5})} = \dfrac{\sqrt{15}}{6}$.

d) Calculamos $\langle A, B\rangle = 0 - 1 + 6 - 3 = 2$, $\|A\|^2 = 4 + 1 + 9 + 1 = 15$, $\|B\|^2 = 0 + 1 + 4 + 9 = 14$. Así

$$\cos\theta = \frac{2}{\sqrt{15}\sqrt{14}} = \frac{2}{\sqrt{210}}$$

6.9. Comprobar:

a) Ley del paralelogramo (Fig. 6-10): $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

b) Forma polar de $\langle u, v\rangle$ (que muestra que el producto interno puede obtenerse de la función norma): $\langle u, v\rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$.

Desarrollamos cada una de las expresiones que siguen para obtener:

$$\|u + v\|^2 = \langle u + v, u + v\rangle = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \qquad [1]$$

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2 \qquad [2]$$

Sumando [1] y [2] llegamos a la ley del paralelogramo *a*). Restando [2] a [1],

$$\|u + v\|^2 - \|u - v\|^2 = 4\langle u, v\rangle$$

Dividimos por 4 para conseguir la forma polar (real) *b*). (La forma polar en el caso complejo es diferente.)

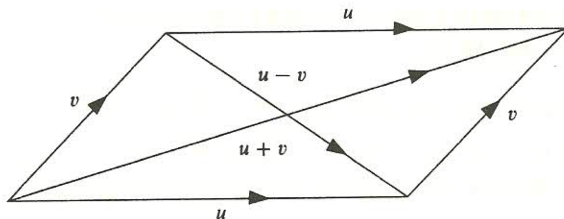


Figura 6-10.

6.10. Demostrar el Teorema 6.1 (Cauchy-Schwarz).

Para cualquier número real t,

$$\langle tu + v, tu + c\rangle = t^2\langle u, u\rangle + 2t\langle u, v\rangle + \langle v, v\rangle = t^2 \| u \|^2 + 2t\langle u, v\rangle + \| v \|^2$$

Sean $a = \|u\|^2$, $b = 2\langle u, v\rangle$ y $c = \|v\|^2$. Como $\|tu + v\|^2 \geqslant 0$, tenemos

$$at^2 + bt + c \geq 0$$

para todo valor de t. Esto significa que el polinomio cuadrático no puede tener dos raíces reales, lo que implica $b^2 - 4ac \leqslant 0$ o $b^2 \leqslant 4ac$. De este modo,

$$4\langle u, v\rangle^2 \leq 4 \| u \|^2 \| v \|^2$$

Dividir por 4 nos conduce a nuestro resultado. (*Nota:* La desigualdad de Cauchy-Schwarz para espacios complejos con producto interno se trata en el Problema 6.53.)

6.11. Demostrar el Teorema 6.2.

Si $v \neq 0$, necesariamente $\langle v, v\rangle > 0$ y por ende $\|v\| = \sqrt{\langle v, v\rangle} > 0$. Si $v = 0$, entonces $\langle 0, 0\rangle = 0$. En consecuencia, $\|0\| = \sqrt{0} = 0$. Siendo así, $[N_1]$ es cierto.

Tenemos $\|kv\|^2 = \langle kv, kv\rangle = k^2\langle v, v\rangle = k^2\|v\|^2$. Tomando la raíz cuadrada de ambos miembros llegamos a $[N_2]$.

Empleando la desigualdad de Cauchy-Schwarz,

$$\| u + v \|^2 = \langle u + v, u + v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle u, v\rangle + \langle v, v\rangle \leq$$
$$\leq \| u \|^2 + 2 \| u \| \| v \| + \| v \|^2 = (\| u \| + \| v \|)^2$$

Tomando la raíz cuadrada de ambos miembros obtenemos $[N_3]$.

6.12. Sean $(a_1, a_2, \ldots)$ y $(b_1, b_2, \ldots)$ cualquier par de vectores en el espacio l_2 del Ejemplo 6.3 *b*). Mostrar que el producto interno está bien definido, es decir, probar que la suma $\sum_{i=1}^{\infty} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots$ converge absolutamente.

Según el Ejemplo 6.4 *a*) (desigualdad de Cauchy-Schwarz),

$$|a_1 b_1| + \cdots + |a_n b_n| \leq \sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2} \leq \sqrt{\sum_{i=1}^{\infty} a_i^2} \sqrt{\sum_{i=1}^{\infty} b_i^2}$$

lo que se verifica para todo n. Así la sucesión (monótona) de sumas $S_n = |a_1 b_1| + \cdots + |a_n b_n|$ es acotada y por tanto converge. Podemos decir, pues, que la suma infinita converge absolutamente.

ORTOGONALIDAD. COMPLEMENTOS ORTOGONALES. CONJUNTOS ORTOGONALES

6.13. Determinar k para que los pares que se citan a continuación sean ortogonales:

a) $u = (1, 2, k, 3)$ y $v = (3, k, 7, -5)$ en $\mathbf{R}^4$.

b) $f(t) = t + k$ y $g(t) = t^2$, donde $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$.

a) Primero hallamos $\langle u, v\rangle = (1, 2, k, 3)\cdot(3, k, 7, -5) = 3 + 2k + 7k - 15 = 9k - 12$. Después tomamos $\langle u, v\rangle = 9k - 12 = 0$ para encontrar $k = \frac{4}{3}$.

b) Primero hallamos

$$\langle f, g\rangle = \int_0^1 (t + k)t^2\, dt = \int_0^1 (t^3 + kt^2)\, dt = \left[\frac{t^4}{4} + \frac{kt^3}{3}\right]_0^1 = \frac{1}{4} + \frac{k}{3}$$

Tomamos $\langle f, g\rangle = \dfrac{1}{4} + \dfrac{k}{3} = 0$ obteniendo $k = -\dfrac{3}{4}$.

6.14. Considérese $u = (0, 1, -2, 5)$ en $\mathbf{R}^4$. Hallar una base para el complemento ortogonal $u^\perp$ de u.

Buscamos todos los vectores (x, y, z, t) de $\mathbf{R}^4$ tales que

$$\langle (x, y, z, t), (0, 1, -2, 5)\rangle = 0 \qquad \text{o} \qquad 0x + y - 2z + 5t = 0$$

Las variables libres son x, z y t. De acuerdo con ello:

1. Tomamos $x = 1$, $z = 0$, $t = 0$ para obtener la solución $w_1 = (1, 0, 0, 0)$.
2. Tomamos $x = 0$, $z = 1$, $t = 0$ para obtener la solución $w_2 = (0, 2, 1, 0)$.
3. Tomamos $x = 0$, $z = 0$, $t = 1$ para obtener la solución $w_3 = (0, -5, 0, 1)$.

Los vectores w_1, w_2, w_3 constituyen una base del espacio solución de la ecuación y por ende una base de $u^\perp$.

6.15. Sea W el subespacio de $\mathbf{R}^5$ generado por $u = (1, 2, 3, -1, 2)$ y $v = (2, 4, 7, 2, -1)$. Hallar una base para el complemento ortogonal $W^\perp$ de W.

Buscamos todos los vectores $w = (x, y, z, s, t)$ tales que

$$\langle w, u\rangle = x + 2y + 3z - s + 2t = 0$$
$$\langle w, v\rangle = 2x + 4y + 7z + 2s - t = 0$$

Despejando x de la segunda ecuación encontramos el sistema equivalente

$$x + 2y + 3z - s + 2t = 0$$
$$z + 4s - 5t = 0$$

Las variables libres son y, s y t. Por eso:

1. Tomamos $y = -1$, $s = 0$, $t = 0$ para obtener la solución $w_1 = (2, -1, 0, 0, 0)$.
2. Tomamos $y = 0$, $s = 1$, $t = 0$ para obtener la solución $w_2 = (13, 0, -4, 1, 0)$.
3. Tomamos $y = 0$, $s = 0$, $t = 1$ para obtener la solución $w_3 = (-17, 0, 5, 0, 1)$.

El conjunto $\{w_1, w_2, w_3\}$ es una base de $W^\perp$.

6.16. Sea $w = (1, 2, 3, 1)$ un vector en $\mathbf{R}^4$. Encontrar una base ortogonal de $w^\perp$.

Hallamos una solución no nula de $x + 2y + 3z + t = 0$; por ejemplo, $v_1 = (0, 0, 1, -3)$. Ahora hallamos una solución no nula del sistema

$$x + 2y + 3z + t = 0 \qquad z - 3t = 0$$

digamos $v_2 = (0, -5, 3, 1)$. Por último, hallamos una solución no nula del sistema

$$x + 2y + 3z + t = 0 \qquad -5y + 3z + t = 0 \qquad z - 3t = 0$$

como puede ser $v_3 = (-14, 2, 3, 1)$. De este modo, v_1, v_2, v_3 forman una base ortogonal de $w^\perp$. (Compárese con el Problema 6.14, donde la base no tenía que ser ortogonal.)

6.17. Supóngase que S consiste en los vectores de $\mathbf{R}^3$:

$$u_1 = (1, 1, 1) \qquad u_2 = (1, 2, -3) \qquad u_3 = (5, -4, -1)$$

a) Probar que S es ortogonal y es una base de $\mathbf{R}^3$.

b) Escribir $v = (1, 5, -7)$ como combinación lineal de u_1, u_2, u_3.

a) Calculamos

$$\langle u_1, u_2\rangle = 1 + 2 - 3 = 0 \qquad \langle u_1, u_3\rangle = 5 - 4 - 1 = 0 \qquad \langle u_2, u_3\rangle = 5 - 8 + 3 = 0$$

Dado que cada producto interno es igual a 0, S es ortogonal y debido a ello es linealmente independiente. Así S es una base de $\mathbf{R}^3$, pues tres vectores linealmente independientes cualesquiera en $\mathbf{R}^3$ constituyen una base del espacio.

b) Sea $v = xu_1 + yu_2 + zu_3$ para escalares desconocidos x, y, z, esto es,

$$(1, 5, -7) = x(1, 1, 1) + y(1, 2, -3) + z(5, -4, -1) \qquad [1]$$

Método 1. Desarrollamos [1] obteniendo el sistema

$$x + y + 5z = 1 \qquad x + 2y - 4z = 5 \qquad x - 3y - z = -7$$

Lo resolvemos llegando a $x = -\frac{1}{3}$, $y = \frac{16}{7}$, $z = -\frac{4}{21}$.

Método 2. (Este método utiliza el hecho de que los vectores de la base son ortogonales, con lo que la aritmética es más sencilla.) Tomamos el producto interno de [1] y u_1 para conseguir

$$(1, 5, -7)\cdot(1, 1, 1) = x(1, 1, 1)\cdot(1, 1, 1) \quad \text{o} \quad -1 = 3x \quad \text{o} \quad x = -\tfrac{1}{3}$$

(Los dos últimos términos desaparecen porque u_1 es ortogonal a u_2 y u_3.) Tomamos el producto interno de [1] y u_2 obteniendo

$$(1, 5, -7)\cdot(1, 2, -3) = y(1, 2, -3)\cdot(1, 2, -3) \quad \text{o} \quad 32 = 14y \quad \text{o} \quad y = \tfrac{16}{7}$$

Tomamos el producto interno de [1] y u_3 llegando a

$$(1, 5, -7)\cdot(5, -4, -1) = z(5, -4, -1)\cdot(5, -4, -1) \quad \text{o} \quad -8 = 42z \quad \text{o} \quad z = -\tfrac{4}{21}$$

En cualquiera de los casos conseguimos $v = (-\frac{1}{3})u_1 + (\frac{16}{7})u_2 - (\frac{4}{21})u_3$.

6.18. Supóngase que S consiste en los vectores de $\mathbf{R}^4$:

$$u_1 = (1, 1, 0, -1) \qquad u_2 = (1, 2, 1, 3) \qquad u_3 = (1, 1, -9, 2) \qquad u_4 = (16, -13, 1, 3)$$

a) Probar que S es ortogonal y es una base de $\mathbf{R}^4$.

b) Determinar las coordenadas de un vector arbitrario $v = (a, b, c, d)$ en $\mathbf{R}^4$ relativas a la base S.

a) Calculamos

$$u_1 \cdot u_2 = 1 + 2 + 0 - 3 = 0 \quad u_1 \cdot u_3 = 1 + 1 + 0 - 2 = 0 \quad u_1 \cdot u_4 = 16 - 13 + 0 - 3 = 0$$
$$u_2 \cdot u_3 = 1 + 2 - 9 + 6 = 0 \quad u_2 \cdot u_4 = 16 - 26 + 1 + 9 = 0 \quad u_3 \cdot u_4 = 16 - 13 - 9 + 6 = 0$$

Siendo así, S es ortogonal y por tanto linealmente independiente. En consecuencia, S es una base de $\mathbf{R}^4$, puesto que cuatro vectores linealmente independientes arbitrarios en $\mathbf{R}^4$ deben serlo.

b) Dado que S es ortogonal, sólo es necesario encontrar los coeficientes de Fourier de v respecto a los vectores de la base, como en el Teorema 6.7. De este modo,

$$k_1 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{a + b - d}{3} \qquad k_3 = \frac{\langle v, u_3\rangle}{\langle u_3, u_3\rangle} = \frac{a + b - 9c + 2d}{87}$$

$$k_2 = \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{a + 2b + c + 3d}{15} \qquad k_4 = \frac{\langle v, u_4\rangle}{\langle u_4, u_4\rangle} = \frac{16a - 13b + c + 3d}{435}$$

son las coordenadas de v respecto a la base S.

6.19. Supóngase que S, S_1 y S_2 son subconjuntos de V. Probar las siguientes aserciones:

a) $S \subseteq S^{\perp\perp}$ *b*) Si $S_1 \subseteq S_2$, necesariamente $S_2^{\perp} \subseteq S_1^{\perp}$ *c*) $S^{\perp} = \text{lin } S^{\perp}$

a) Sea $w \in S$. Entonces $\langle w, v\rangle = 0$ para todo $v \in S^{\perp}$; por tanto, $w \in S^{\perp\perp}$. De acuerdo con esto, $S \subseteq S^{\perp\perp}$.

b) Sea $w \in S_2^{\perp}$. En tal caso, $\langle w, v\rangle = 0$ para todo $v \in S_2$. Como $S_1 \subseteq S_2$, $\langle w, v\rangle = 0$ para todo $v \in S_1$. Luego $w \in S_1^{\perp}$ y por consiguiente $S_2^{\perp} \subseteq S_1^{\perp}$.

c) Como $S \subseteq \text{lin } S$, tenemos $\text{lin } S^{\perp} \subseteq S^{\perp}$. Supongamos que $u \in S^{\perp}$ y $v \in \text{lin } S^{\perp}$. Existen, pues, vectores $w_1, w_2, \ldots, w_k$ en S tales que

$$v = a_1 w_1 + a_2 w_2 + \cdots + a_k w_k$$

Utilizando $u \in S^{\perp}$,

$$\langle u, v\rangle = \langle u, a_1 w_1 + a_2 w_2 + \cdots + a_k w_k\rangle = a_1\langle u, w_1\rangle + a_2\langle u, w_2\rangle + \cdots + a_k\langle u, w_k\rangle =$$
$$= a_1 \cdot 0 + a_2 \cdot 0 + \cdots + a_k \cdot 0 = 0$$

Así $u \in \text{lin } S^{\perp}$, y en consecuencia $S^{\perp} \subseteq \text{lin } S^{\perp}$. Ambas inclusiones conducen a $S^{\perp} = \text{lin } S^{\perp}$.

6.20. Demostrar el Teorema 6.5.

Supongamos $S = \{u_1, u_2, \ldots, u_r\}$ y

$$a_1 u_1 + a_2 u_2 + \cdots + a_r u_r = 0 \qquad [1]$$

Tomando el producto interno de [1] y u_1,

$$0 = \langle 0, u_1\rangle = \langle a_1u_1 + a_2u_2 + \cdots + a_ru_r, u_1\rangle =$$
$$= a_1\langle u_1, u_1\rangle + a_2\langle u_2, u_1\rangle + \cdots + a_r\langle u_r, u_1\rangle =$$
$$= a_1\langle u_1, u_1\rangle + a_2 \cdot 0 + \cdots + a_r \cdot 0 = a_1\langle u_1, u_1\rangle$$

Puesto que $u_1 \neq 0$, $\langle u_1, u_1\rangle \neq 0$, de forma que $a_1 = 0$. Similarmente, para $i = 2, \ldots, r$, tomando el producto interno de [1] y u_i,

$$0 = \langle 0, u_i\rangle = \langle a_1u_1 + \cdots + a_ru_r, u_i\rangle =$$
$$= a_1\langle u_1, u_i\rangle + \cdots + a_i\langle u_i, u_i\rangle + \cdots + a_r\langle u_r, u_i\rangle = a_i\langle u_i, u_i\rangle$$

Pero $\langle u_i, u_i\rangle \neq 0$ y por tanto $a_i = 0$. Así pues, S es linealmente independiente.

6.21. Demostrar el Teorema 6.6 (Pitágoras).

Desarrollando el producto interno,

$$\|u_1 + u_2 + \cdots + u_r\|^2 = \langle u_1 + u_2 + \cdots + u_r, u_1 + u_2 + \cdots + u_r\rangle =$$
$$= \langle u_1, u_1\rangle + \langle u_2, u_2\rangle + \cdots + \langle u_r, u_r\rangle + \sum_{i \neq j}\langle u_i, u_j\rangle$$

El teorema se deduce del hecho de que $\langle u_i, u_i\rangle = \|u_i\|^2$ y $\langle u_i, u_j\rangle = 0$ para $i \neq j$.

6.22. Demostrar el Teorema 6.7.

Supongamos $v = k_1u_1 + k_2u_2 + \cdots + k_nu_n$. Tomando el producto interno de ambos miembros con u_1

$$\langle v, u_1\rangle = \langle k_1u_1 + k_2u_2 + \cdots + k_nu_n, u_1\rangle =$$
$$= k_1\langle u_1, u_1\rangle + k_2\langle u_2, u_1\rangle + \cdots + k_n\langle u_n, u_1\rangle =$$
$$= k_1\langle u_1, u_1\rangle + k_2 \cdot 0 + \cdots + k_n \cdot 0 = k_1\langle u_1, u_1\rangle$$

De este modo, $k_1 = \langle v, u_1\rangle / \langle u_1, u_1\rangle$. Análogamente, para $i = 2, \ldots, n$,

$$\langle v, u_i\rangle = \langle k_1u_1 + k_2u_2 + \cdots + k_nu_n, u_i\rangle =$$
$$= k_1\langle u_1, u_i\rangle + k_2\langle u_2, u_i\rangle + \cdots + k_n\langle u_n, u_i\rangle =$$
$$= k_1 \cdot 0 + \cdots + k_i\langle u_i, u_i\rangle + \cdots + k_n \cdot 0 = k_i\langle u_i, u_i\rangle$$

De este modo, $k_i = \langle v, u_i\rangle / \langle u_i, u_i\rangle$. Sustituyendo la expresión de cada k_i en la ecuación $v = k_1u_1 + \cdots + k_nu_n$ conseguimos el resultado deseado.

6.23. Supóngase que $E = \{e_1, e_2, \ldots, e_n\}$ es una base ortonormal de V. Demostrar las afirmaciones:

a) Para todo $u \in V$ tenemos $u = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \cdots + \langle u, e_n\rangle e_n$.

b) $\langle a_1e_1 + \cdots + a_ne_n, b_1e_1 + \cdots + b_ne_n\rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

c) Para todo par de vectores $u, v \in V$ tenemos $\langle u, v\rangle = \langle u, e_1\rangle\langle v, e_1\rangle + \cdots + \langle u, e_n\rangle\langle v, e_n\rangle$.

a) Sea $u = k_1e_1 + k_2e_2 + \cdots + k_ne_n$. Tomando el producto interno de u y e_1,

$$\langle u, e_1\rangle = \langle k_1e_1 + k_2e_2 + \cdots + k_ne_n, e_1\rangle =$$
$$= k_1\langle e_1, e_1\rangle + k_2\langle e_2, e_1\rangle + \cdots + k_n\langle e_n, e_1\rangle =$$
$$= k_1 \cdot 1 + k_2 \cdot 0 + \cdots + k_n \cdot 0 = k_1$$

De forma similar, para $i = 2, \ldots, n$,

$$\begin{aligned}\langle u, e_i\rangle &= \langle k_1 e_1 + \cdots + k_i e_i + \cdots + k_n e_n, e_i\rangle = \\ &= k_1\langle e_1, e_i\rangle + \cdots + k_i\langle e_i, e_i\rangle + \cdots + k_n\langle e_n, e_i\rangle = \\ &= k_1 \cdot 0 + \cdots + k_i \cdot 1 + \cdots + k_n \cdot 0 = k_i\end{aligned}$$

Sustituyendo k_i por $\langle u, e_i\rangle$ en la ecuación $u = k_1 e_1 + \cdots + k_n e_n$ llegamos al resultado buscado.

b) Tenemos

$$\left\langle \sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{n} b_j e_j \right\rangle = \sum_{i,j=1}^{n} a_j b_j \langle e_i, e_j\rangle = \sum_{i=1}^{n} a_i b_i \langle e_i, e_i\rangle + \sum_{i \neq 1} a_i b_j \langle e_i, e_j\rangle$$

Pero $\langle e_i, e_j\rangle = 0$ para $i \neq j$ y $\langle e_i, e_j\rangle = 1$ para $i = j$; por tanto, como se pedía,

$$\left\langle \sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{n} b_j e_j \right\rangle = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

c) De acuerdo con la parte *a*),

$$u = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n \qquad \text{y} \qquad v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$$

Y por *b*),

$$\langle u, v\rangle = \langle u, e_1\rangle\langle v, e_1\rangle + \langle u, e_2\rangle\langle v, e_2\rangle + \cdots + \langle u, e_n\rangle\langle v, e_n\rangle$$

PROYECCIONES. ALGORITMO DE GRAM-SCHMIDT. APLICACIONES

6.24. Supóngase $w \neq 0$. Sea v cualquier vector en V. Mostrar que

$$c = \frac{\langle v, w\rangle}{\langle w, w\rangle} = \frac{\langle v, w\rangle}{\|w\|^2}$$

es el único escalar tal que $v' = v - cw$ es ortogonal a w.

Para que v' sea ortogonal a w debe cumplirse

$$\langle v - cw, w\rangle = 0 \qquad \text{o} \qquad \langle v, w\rangle - c\langle w, w\rangle = 0 \qquad \text{o} \qquad \langle v, w\rangle = c\langle w, w\rangle$$

Siendo así, $c = \langle v, w\rangle/\langle w, w\rangle$. Recíprocamente, supongamos $c = \langle v, w\rangle/\langle w, w\rangle$. En tal caso,

$$\langle v - cw, w\rangle = \langle v, w\rangle - c\langle w, w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, w\rangle = 0$$

6.25. Hallar el coeficiente de Fourier c y la proyección de $v = (1, -2, 3, -4)$ a lo largo de $w = (1, 2, 1, 2)$ en $\mathbf{R}^4$.

Calculamos $\langle v, w\rangle = 1 - 4 + 3 - 8 = -8$ y $\|w\|^2 = 1 + 4 + 1 + 4 = 10$. Entonces $c = -\frac{8}{10} = -\frac{4}{5}$ y proy $(v, w) = cw = (-\frac{4}{5}, -\frac{8}{5}, -\frac{4}{5}, -\frac{8}{5})$.

6.26. Encontrar una base ortonormal para el subespacio U de $\mathbf{R}^4$ generado por los vectores

$$v_1 = (1, 1, 1, 1) \qquad v_2 = (1, 1, 2, 4) \qquad v_3 = (1, 2, -4, -3)$$

Primero hallamos una base ortogonal de U empleando el algoritmo de Gram-Schmidt. Iniciamos el algoritmo tomando $w_1 = u_1 = (1, 1, 1, 1)$. A continuación calculamos

$$v_2 - \frac{\langle v_2, w_1 \rangle}{\| w_1 \|^2} w_1 = (1, 1, 2, 4) - \frac{8}{4}(1, 1, 1, 1) = (-1, -1, 0, 2)$$

Tomamos $w_2 = (-1, -1, 0, 2)$. Después calculamos

$$v_3 - \frac{\langle v_3, w_1 \rangle}{\| w_1 \|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\| w_2 \|^2} w_2 = (1, 2, -4, -3) - \frac{-4}{4}(1, 1, 1, 1) - \frac{-9}{6}(-1, -1, 0, 2) =$$
$$= (\tfrac{1}{2}, \tfrac{3}{2}, -3, 1)$$

Eliminamos los denominadores para obtener $w_3 = (1, 3, -6, 2)$. Por último, normalizamos la base ortogonal constituida por w_1, w_2, w_3. Como $\|w_1\|^2 = 4$, $\|w_2\|^2 = 6$ y $\|w_3\|^2 = 50$, los siguientes vectores forman una base ortonormal de U:

$$u_1 = \frac{1}{2}(1, 1, 1, 1) \qquad u_2 = \frac{1}{\sqrt{6}}(-1, -1, 0, 2) \qquad u_3 = \frac{1}{5\sqrt{2}}(1, 3, -6, 2)$$

6.27. Sea V el espacio vectorial de los polinomios $f(t)$ con producto interno $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. Aplicar el algoritmo de Gram-Schmidt al conjunto $\{1, t, t^2\}$ para conseguir un conjunto ortogonal $\{f_0, f_1, f_2\}$, con coeficientes enteros.

Primero tomamos $f_0 = 1$. A continuación hallamos

$$t - \frac{\langle t, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 = t - \frac{\frac{1}{2}}{1} \cdot 1 = t - \frac{1}{2}$$

Eliminamos los denominadores obteniendo $f_1 = 2t - 1$. Después hallamos

$$t^2 - \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle t^2, 2t-1 \rangle}{\langle 2t-1, 2t-1 \rangle} \cdot (2t-1) = t^2 - \frac{\frac{1}{3}}{1} \cdot 1 - \frac{\frac{1}{6}}{\frac{1}{3}} \cdot (2t-1) = t^2 - t + \frac{1}{6}$$

Eliminamos los denominadores para llegar a $f_2 = 6t^2 - 6t + 1$. De este modo, $\{1, 2t - 1, 6t^2 - 6t + 1\}$ es el conjunto ortogonal requerido.

6.28. Supóngase $v = (1, 3, 5, 7)$. Determinar la proyección de v sobre W (o encontrar un vector $w \in W$ que minimice $\|v - w\|$), siendo W el espacio de $\mathbf{R}^4$ generador por:

a) $u_1 = (1, 1, 1, 1)$ y $u_2 = (1, -3, 4, -2)$.

b) $v_1 = (1, 1, 1, 1)$ y $v_2 = (1, 2, 3, 2)$.

a) Dado que u_1 y u_2 son ortogonales, sólo necesitamos calcular los coeficientes de Fourier:

$$c_1 = \frac{\langle v, u_1 \rangle}{\| u_1 \|^2} = \frac{1+3+5+7}{1+1+1+1} = \frac{16}{4} = 4$$

$$c_2 = \frac{\langle v, u_2 \rangle}{\| u_2 \|^2} = \frac{1-9+20-14}{1+9+16+4} = \frac{-2}{30} = \frac{-1}{15}$$

Entonces

$$w = \text{proy}\,(v, W) = c_1 u_1 + c_2 u_2 = 4(1, 1, 1, 1) - \tfrac{1}{15}(1, -3, 4, -2) = (\tfrac{59}{15}, \tfrac{63}{5}, \tfrac{56}{15}, \tfrac{62}{15})$$

b) Como v_1 y v_2 no son ortogonales, comenzamos por aplicar el algoritmo de Gram-Schmidt para hallar una base ortogonal de W. Tomemos $w_1 = v_1 = (1, 1, 1, 1)$. Después hallamos

$$v_2 - \frac{\langle v_2, w_1 \rangle}{w_1^2} w_1 = (1, 2, 3, 2) - \frac{8}{4}(1, 1, 1, 1) = (-1, 0, 1, 0)$$

Tomemos $w_2 = (-1, 0, 1, 0)$. Ahora calculamos

$$c_1 = \frac{\langle v, w_1 \rangle}{\| w_1 \|^2} = \frac{1+3+5+7}{1+1+1+1} = \frac{16}{4} = 4 \qquad \text{y} \qquad c_2 = \frac{\langle v, w_2 \rangle}{\| w_2 \|^2} = \frac{-1+0+5+0}{1+0+1+0} = \frac{-6}{2} = -3$$

Por tanto, $w = \text{proy}\,(v, W) = c_1 w_1 + c_2 w_2 = 4(1, 1, 1, 1) - 3(-1, 0, 1, 0) = (7, 4, 1, 4)$.

6.29. Supóngase que w_1 y w_2 son vectores ortogonales no nulos. Sea v cualquier vector en V. Encontrar c_1 y c_2 de forma que v' sea ortogonal a w_1 y w_2, donde $v' = v - cw_1 - cw_2$.

Si v' es ortogonal a w_1,

$$0 = \langle v - c_1 w_1 - c_2 w_2, w_1 \rangle = \langle v, w_1 \rangle - c_1 \langle w_1, w_1 \rangle - c_2 \langle w_2, w_1 \rangle =$$
$$= \langle v, w_1 \rangle - c_1 \langle w_1, w_1 \rangle - c_2 0 = \langle v, w_1 \rangle - c_1 \langle w_1, w_1 \rangle$$

Siendo así, $c_1 = \langle v, w_1 \rangle / \langle w_1, w_1 \rangle$. (Esto es, c_1 es la componente de v a lo largo de w_1.) Análogamente, si v' es ortogonal a w_2,

$$0 = \langle v - c_1 w_1 - c_2 w_2, w_2 \rangle = \langle v, w_2 \rangle - c_2 \langle w_2, w_2 \rangle$$

Así pues, $c_2 = \langle v, w_2 \rangle / \langle w_2, w_2 \rangle$. (Esto es, c_2 es la componente de v a lo largo de w_2.)

6.30. Demostrar el Teorema 6.8.

Para $i = 1, 2, \ldots, r$ y usando $\langle w_i, w_j \rangle = 0$ para $i \neq j$,

$$\langle v - c_1 w_1 - c_2 w_2 - \cdots - c_r w_r, w_i \rangle = \langle v, w_i \rangle - c_1 \langle w_1, w_i \rangle - \cdots - c_i \langle w_i w_i \rangle - \cdots - c_r \langle w_r, w_i \rangle =$$
$$= \langle v, w_i \rangle - c_1 \cdot 0 - \cdots - c_i \langle w_i, w_i \rangle - \cdots - c_r \cdot 0 =$$
$$= \langle v, w_i \rangle - c_i \langle w_i, w_i \rangle = \langle v, w_i \rangle - \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_i \rangle =$$
$$= 0$$

De este modo queda probado el teorema.

6.31. Demostrar el Teorema 6.9.

Por el Teorema 6.8, $v - \sum c_k w_k$ es ortogonal a todo w_i y por ende ortogonal a cualquier combinación lineal de $w_1, w_2, \ldots, w_r$. Por tanto, usando el teorema de Pitágoras y sumando desde $k = 1$ hasta r,

$$\left\| v - \sum a_k w_k \right\|^2 = \left\| v - \sum c_k w_k + \sum (c_k - a_k) w_k \right\|^2 = \left\| v - \sum c_k w_k \right\|^2 + \left\| \sum (c_k - a_k) w_k \right\|^2 \geq$$
$$\geq \left\| v - \sum c_k w_k \right\|^2$$

La raíz cuadrada de ambos miembros nos proporciona nuestro teorema.

6.32. Demostrar el Teorema 6.10 (desigualdad de Bessel).

Nótese que $c_i = \langle v, e_i \rangle$ puesto que $\|e_i\| = 1$. Entonces, utilizando $\langle e_i, e_j \rangle = 0$ para $i \neq j$ y sumando desde $k = 1$ hasta r,

$$0 \le \left\langle v - \sum c_k e_k, v - \sum c_k, e_k \right\rangle = \langle v, v \rangle - 2\left\langle v, \sum c_k e_k \right\rangle + \sum c_k^2 =$$

$$= \langle v, v \rangle - \sum 2c_k \langle v, e_k \rangle + \sum c_k^2 = \langle v, v \rangle - \sum 2c_k^2 + \sum c_k^2 =$$

$$= \langle v, v \rangle - \sum c_k^2$$

Esto nos conduce a nuestra desigualdad.

6.33. Demostrar el Teorema 6.11.

La demostración hace uso del algoritmo de Gram-Schmidt y de las Notas 1 y 2 que le siguen (Sección 6.6). Concretamente, aplicamos el algoritmo a $\{v_i\}$ para conseguir una base ortogonal $\{w_1, \dots, w_n\}$, y después normalizamos $\{w_i\}$ para obtener una base ortonormal $\{u_i\}$ de V. El algoritmo garantiza que cada w_k es combinación lineal de $v_1, \dots, v_k$ y por consiguiente también lo es cada u_k.

6.34. Demostrar el Teorema 6.12.

Extendemos S a una base $S' = \{w_1, \dots, w_r, v_{r+1} \dots, v_n\}$ de V. Aplicando el algoritmo de Gram-Schmidt a S' obtenemos primero $w_1, w_2, \dots, w_r$, porque S es ortogonal, y luego vectores $w_{r+1}, \dots, w_n$, siendo $\{w_1, w_2, \dots, w_n\}$ una base ortogonal de V. Queda así probado el teorema.

6.35. Demostrar el Teorema 6.4.

Según el Teorema 6.11, existe una base ortogonal $\{u_1, \dots, u_r\}$ de W que, de acuerdo con el Teorema 6.12, podemos extender a una base ortogonal $\{u_1, u_2, \dots, u_n\}$ de V. Por tanto, $u_{r+1}, \dots, u_n \in W^\perp$. Si $v \in V$,

$$v = a_1 u_1 + \cdots + a_n u_n \qquad \text{donde} \qquad a_1 u_1 + \cdots + a_r u_r \in W, \qquad a_{r+1} u_{r+1} + \cdots + a_n u_n \in W^\perp$$

De acuerdo con ello, $V = W + W^\perp$.

Por otra parte, si $w \in W \cap W^\perp$, necesariamente $\langle w, w \rangle = 0$. Esto nos conduce a $w = 0$; de aquí que sea $W \cap W^\perp = \{0\}$.

Las dos condiciones, $V = W + W^\perp$ y $W \cap W^\perp = \{0\}$, llevan al resultado deseado $V = W \oplus W^\perp$.

Adviértase que hemos demostrado el teorema únicamente en el caso en que V tiene dimensión finita. Hacemos notar que el teorema es igualmente válido para espacios de dimensión arbitraria.

6.36. Supóngase que W es un subespacio de un espacio de dimensión finita V. Probar que $W = W^{\perp\perp}$.

Según el Teorema 6.4, $V = W \oplus W^\perp$ y también $V = W^\perp \oplus W^{\perp\perp}$. Por este motivo,

$$\dim W = \dim V - \dim W^\perp \qquad \text{y} \qquad \dim W^{\perp\perp} = \dim V - \dim W^\perp$$

Esto proporciona $\dim W = \dim W^{\perp\perp}$. Pero $W \subseteq W^{\perp\perp}$ [Problema 6.19 *a*)], luego $W = W^{\perp\perp}$, como se requería.

PRODUCTOS INTERNOS Y MATRICES DEFINIDAS POSITIVAS

6.37. Determinar si A es o no definida positiva, donde $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

Empleamos el algoritmo de diagonalización de la Sección 4.11 para transformar A en una matriz (congruente) diagonal.

Efectuamos $-R_1 + R_3 \to R_3$ y $-C_1 + C_3 \to C_3$ y después $-2R_2 + R_3 \to R_3$ y $-2C_2 + C_3 \to C_3$ para obtener

$$A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Hay una entrada negativa, -2, en la matriz diagonal, luego A no es definida positiva.

6.38. Hallar la matriz A que representa el producto interno usual en $\mathbf{R}^2$ respecto a cada una de las siguientes bases:

a) $\{v_1 = (1, 3), v_2 = (2, 5)\}$, *b*) $\{w_1 = (1, 2), w_2 = (4, -2)\}$.

a) Calculamos $\langle v_1, v_1 \rangle = 1 + 9 = 10$, $\langle v_1, v_2 \rangle = 2 + 15 = 17$, $\langle v_2, v_2 \rangle = 4 + 25 = 29$. Así $A = \begin{pmatrix} 10 & 17 \\ 17 & 29 \end{pmatrix}$.

b) Calculamos $\langle w_1, w_1 \rangle = 1 + 4 = 5$, $\langle w_1, w_2 \rangle = 4 - 4 = 0$, $\langle w_2, w_2 \rangle = 16 + 4 = 20$. De este modo, $A = \begin{pmatrix} 5 & 0 \\ 0 & 20 \end{pmatrix}$.

(Siendo ortogonales los vectores de la base, la matriz A es diagonal.)

6.39. Considérese el espacio vectorial V de los polinomios $f(t)$ de grado $\leqslant 2$ con producto interno $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)dt$.

a) Calcular $\langle f, g \rangle$ para $f(t) = t + 2$ y $g(t) = t^2 - 3t + 4$.

b) Hallar la matriz A del producto interno respecto a la base $\{1, t, t^2\}$ de V.

c) Comprobar el Teorema 6.14 mostrando que $\langle f, g \rangle = [f]^T A[g]$ respecto a la base $\{1, t, t^2\}$.

a) $$\langle f, g \rangle = \int_{-1}^{1} (t + 2)(t^2 - 3t + 4)\, dt = \int_{-1}^{1} (t^3 - t^2 - 2t + 8)\, dt = \left[\frac{t^4}{4} - \frac{t^3}{3} - t^2 + 8t\right]_{-1}^{1} = \frac{46}{3}.$$

b) Aquí utilizamos el hecho de que, si $r + s = n$,

$$\langle t^r, t^s \rangle = \int_{-1}^{1} t^n\, dt = \left[\frac{t^{n+1}}{n+1}\right]_{-1}^{1} = \begin{cases} 2/(n+1) & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar} \end{cases}$$

Entonces $\langle 1, 1 \rangle = 2$, $\langle 1, t \rangle = 0$, $\langle 1, t^2 \rangle = \frac{2}{3}$, $\langle t, t \rangle = \frac{2}{3}$, $\langle t, t^2 \rangle = 0$, $\langle t^2, t^2 \rangle = \frac{2}{5}$. Siendo así,

$$A = \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix}$$

c) Tenemos $[f]^T = (2, 1, 0)$ y $[g]^T = (4, -3, 1)$ respecto a la base dada. En tal caso,

$$[f]^T A[g] = (2, 1, 0)\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix}\begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix} = (4, \tfrac{2}{3}, \tfrac{4}{3})\begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix} = \tfrac{46}{3} = \langle f, g\rangle$$

6.40. Demostrar el Teorema 6.13.

Para vectores arbitrarios u_1, u_2 y v,

$$\langle v_1 + u_2, v\rangle = (u_1 + u_2)^T Av = (u_1^T + u_2^T)Av = u_1^T Av + u_2^T Av = \langle u_1, v\rangle + \langle u_2, v\rangle$$

y, para un escalar k y dos vectores u, v cualesquiera,

$$\langle ku, v\rangle = (ku)^T Av = ku^T Av = k\langle u, v\rangle$$

De este modo se verifica $[I_1]$.

Como $u^T Au$ es un escalar, $(u^T Av)^T = u^T Av$. Además, $A^T = A$, ya que A es simétrica. Por tanto,

$$\langle u, v\rangle = u^T Av = (u^T Av)^T = v^T A^T t^{TT} = v^T Au = \langle v, u\rangle$$

Así pues, se satisface $[I_2]$.

Finalmente, al ser A definida positiva, $X^T AX > 0$ para todo $X \in \mathbf{R}^n$ no nulo. Para todo vector no nulo tendremos, pues, $\langle v, v\rangle = v^T Av > 0$. Además, $\langle 0, 0\rangle = 0^T A0 = 0$. Siendo así, se verifica $[I_3]$. Con todo ello, la función $\langle u, v\rangle = u^T Av$ es un producto interno.

6.41. Demostrar el Teorema 6.14.

Supongamos $S = \{w_1, w_2, \ldots, w_n\}$ y $A = (k_{ij})$. En consecuencia, $k_{ij} = \langle w_i, w_j\rangle$. Supongamos

$$u = a_1 w_1 + a_2 w_2 + \cdots + a_n w_n \qquad \text{y} \qquad v = b_1 w_1 + b_2 w_2 + \cdots + b_n w_n$$

Entonces

$$\langle u, v\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \langle w_i, w_j\rangle \qquad [1]$$

Por otra parte,

$$[u]^T A[v] = (a_1, a_2, \ldots, a_n)\begin{pmatrix} k_{11} & k_{22} & \ldots & k_{1n} \\ k_{21} & k_{22} & \ldots & k_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ k_{n1} & k_{n2} & \ldots & k_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} =$$

$$= \left(\sum_{i=1}^{n} a_i k_{i1}, \sum_{i=1}^{n} a_i k_{i2}, \ldots, \sum_{i=1}^{n} a_i k_{in}\right)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_i b_j k_{ij} \qquad [2]$$

Dado que $k_{ij} = \langle w_i, w_j\rangle$, las sumas finales de [1] y [2] son iguales, de donde $\langle u, v\rangle = [u]^T A[v]$.

6.42. Demostrar el Teorema 6.15.

Puesto que $\langle w_i, w_j\rangle = \langle w_j, w_i\rangle$ para todo par de vectores de la base w_i y w_j, la matriz A es simétrica. Sea X cualquier vector no nulo en $\mathbf{R}^n$. Entonces $[u] = X$ para algún vector no nulo $u \in V$. Haciendo uso del Teorema 6.14 tenemos $X^T AX = [u]^T A[u] = \langle u, u\rangle > 0$, de modo que A es definida positiva.

PRODUCTOS INTERNOS Y MATRICES ORTOGONALES

6.43. Encontrar una matriz ortogonal P cuya primera fila sea $u_1 = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

Empezamos por hallar un vector no nulo $w_2 = (x, y, z)$ que sea ortogonal a u_1, es decir, para el que

$$0 = \langle u_1, w_2 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{o} \qquad x + 2y + 2z = 0$$

Una de estas soluciones es $w_2 = (0, 1, -1)$. Normalizamos w_2 para conseguir la segunda fila de P, esto es, $u_2 = (0, 1/\sqrt{2}, -1/\sqrt{2})$.

Seguidamente hallamos un vector no nulo $w_3 = (x, y, z)$ que sea ortogonal tanto a u_1 como a u_2, o sea, para el que

$$0 = \langle u_1, w_3 \rangle = \frac{x}{3} + \frac{2y}{3} + \frac{2z}{3} = 0 \qquad \text{o} \qquad x + 2y + 2z = 0$$

$$0 = \langle u_2, w_3 \rangle = \frac{y}{\sqrt{2}} - \frac{z}{\sqrt{2}} = 0 \qquad \text{o} \qquad y - z = 0$$

Tomamos $z = -1$ y encontramos la solución $w_3 = (4, -1, -1)$. Normalizando w_3 obtenemos la tercera fila de P, esto es, $u_3 = (4/\sqrt{18}, -1/\sqrt{18}, -1/\sqrt{18})$. Por tanto,

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 4/3\sqrt{2} & -1/3\sqrt{2} & -1/3\sqrt{2} \end{pmatrix}$$

Subrayamos que la matriz anterior no es única.

6.44. Sea $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & 4 \\ 7 & -5 & 2 \end{pmatrix}$. Determinar si: *a*) las filas de A son ortogonales, *b*) A es una matriz ortogonal, *c*) las columnas de A son ortogonales.

a) Sí, porque

$$(1, 1, -1) \cdot (1, 3, 4) = 1 + 3 - 4 = 0 \qquad (1, 1, -1) \cdot (7, -5, 2) = 7 - 5 - 2 = 0$$
$$(1, 3, 4) \cdot (7, -5, 2) = 7 - 15 + 8 = 0$$

b) No, pues las filas de A no son vectores unitarios, por ejemplo, $(1, 1, -1)^2 = 1 + 1 + 1 = 3$.

c) No, por ejemplo, $(1, 1, 7) \cdot (1, 3, -5) = 1 + 3 - 35 = -31 \neq 0$.

6.45. Sea B la matriz obtenida normalizando cada fila de la matriz A del Problema 6.44. *a*) Hallar B. *b*) ¿Es B una matriz ortogonal? *c*) ¿Son las columnas de B ortogonales?

a) Tenemos

$$\|(1, 1, -1)\|^2 = 1 + 1 + 1 = 3 \qquad \|(1, 3, 4)\|^2 = 1 + 9 + 16 = 26$$
$$\|(7, -5, 2)\|^2 = 49 + 25 + 4 = 78$$

Así pues,

$$B = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 1/\sqrt{26} & 3/\sqrt{26} & 4/\sqrt{26} \\ 7/\sqrt{78} & -5/\sqrt{78} & 2/\sqrt{78} \end{pmatrix}$$

b) Sí, puesto que las filas de B continúan siendo ortogonales y ahora son vectores unitarios.

c) Sí, ya que si las filas de B forman un conjunto ortonormal de vectores, por el Teorema 6.15, sus columnas deben formar, automáticamente, un conjunto ortonormal.

6.46. Demostrar las afirmaciones que se hacen a continuación:

a) P es ortogonal si y sólo si lo es P^T.

b) Si P es ortogonal, necesariamente lo es P^{-1}.

c) Si P y Q son ortogonales, PQ es ortogonal.

a) Sabemos que $(P^T)^T = P$. De este modo, P es ortogonal si y sólo si $PP^T = I$, si y sólo si $P^{TT}P^T = I$, si y sólo si P^T es ortogonal.

b) Tenemos $P^T = P^{-1}$, por ser P ortogonal. Así, de acuerdo con *a*), P^{-1} es ortogonal.

c) Tenemos $P^T = P^{-1}$ y $Q^T = Q^{-1}$, de modo que

$$(PQ)(PQ)^T = PQQ^TP^T = PQQ^{-1}P^{-1} = I$$

Siendo así, $(PQ)^T = (PQ)^{-1}$ y PQ es, pues, ortogonal.

6.47. Supóngase que P es una matriz ortogonal. Mostrar que:

a) $\langle Pu, Pv\rangle = \langle u, v\rangle$ para vectores cualesquiera $u, v \in V$; *b*) $\|Pu\| = \|u\|$ para todo $u \in V$.

a) Usando $P^TP = I$,

$$\langle Pu, Pv\rangle = (Pu)^T(Pv) = u^TP^TPv = u^Tv = \langle u, v\rangle$$

b) Usando $P^TP = I$,

$$\|Pu\|^2 = \langle Pu, Pu\rangle = u^TP^TPu = u^Tu = \langle u, u\rangle = \|u\|^2$$

Tomar la raíz cuadrada de ambos miembros nos conduce a nuestro resultado.

6.48. Demostrar el Teorema 6.17.

Supongamos

$$e'_i = b_{i1}e_1 + b_{i2}e_2 + \cdots + b_{in}e_n \qquad i = 1, \ldots, n \qquad [1]$$

Utilizando el Problema 6.23 *b*) y el hecho de que E' es ortonormal,

$$\delta_{ij} = \langle e'_i, e'_j\rangle = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \qquad [2]$$

Sea $B = (b_{ij})$ la matriz de coeficientes en [1]. (Entonces $P = B^T$.) Supongamos $BB^T = (c_{ij})$,

$$c_{ij} = b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn} \qquad [3]$$

Por [2] y [3], $c_{ij} = \delta_{ij}$. De este modo, $BB^T = I$. En consecuencia, B es ortogonal y por consiguiente $P = B^T$ lo es también.

6.49. Demostrar el Teorema 6.18.

Como $\{e_i\}$ es ortonormal, obtenemos, según el Problema 6.23 *b*),

$$\langle e'_i, e'_j \rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \langle C_i, C_j \rangle$$

donde C_i denota la columna i-ésima de la matriz ortogonal $P = (a_{ij})$. Siendo P ortogonal, sus columnas constituyen un conjunto ortonormal, lo que implica $\langle e'_i, e'_j \rangle = \langle C_i, C_j \rangle = \delta_{ij}$. Por tanto, $\{e'_i\}$ es una base ortonormal.

ESPACIOS COMPLEJOS CON PRODUCTO INTERNO

6.50. Sea V un espacio complejo con producto interno. Comprobar que se verifica la relación

$$\langle u, av_1 + bv_2 \rangle = \bar{a}\langle u, v_1 \rangle + \bar{b}\langle u, v_2 \rangle$$

Utilizando $[I_2^*]$, $[I_1^*]$ y luego $[I_2^*]$ encontramos

$$\langle u, av_1 + bv_2 \rangle = \overline{\langle av_1 + bv_2, u \rangle} = \overline{a\langle v_1, u \rangle + b\langle v_2, u \rangle} = \bar{a}\overline{\langle v_1, u \rangle} + \bar{b}\overline{\langle v_2, u \rangle} = \bar{a}\langle u, v_1 \rangle + \bar{b}\langle u, v_2 \rangle$$

6.51. Supóngase que $\langle u, v \rangle = 3 + 2i$ en un espacio complejo con producto interno V. Calcular:

a) $\langle (2 - 4i)u, v \rangle$; *b*) $\langle u, (4 + 3i)v \rangle$; *c*) $\langle (3 - 6i)u, (5 - 2i)v \rangle$

a) $\langle (2 - 4i)u, v \rangle = (2 - 4i)\langle u, v \rangle = (2 - 4i)(3 + 2i) = 14 - 18i$.

b) $\langle u, (4 + 3i)v \rangle = \overline{(4 + 3i)}\langle u, v \rangle = (4 - 3i)(3 + 2i) = 18 - i$.

c) $\langle (3 - 6i)u, (5 - 2i)v \rangle = (3 - 6i)\overline{(5 - 2i)}\langle u, v \rangle = (3 - 6i)(5 + 2i)(3 + 2i) = 137 - 30i$.

6.52. Hallar el coeficiente de Fourier (o componente) c y la proyección cw de $v = (3 + 4i, 2 - 3i)$ a lo largo de $w = (5 + i, 2i)$ en $\mathbf{C}^2$.

Recuérdese que $c = \langle v, w \rangle / \langle w, w \rangle$. Calculamos

$$\langle v, w \rangle = (3 + 4i)\overline{(5 + i)} + (2 - 3i)\overline{(2i)} = (3 + 4i)(5 - i) + (2 - 3i)(-2i) =$$
$$= 19 + 17i - 6 - 4i = 13 + 13i$$
$$\langle w, w \rangle = 25 + 1 + 4 = 30$$

Así $c = (13 + 13i)/30 = \frac{13}{30} + 13i/30$. En consecuencia,

$$\text{proy } (v, w) = cw = (\tfrac{26}{15} + 39i/15, -\tfrac{13}{15} + i/15)$$

6.53. Demostrar el Teorema 6.19 (Cauchy-Schwarz).

Si $v = 0$, la desigualdad se reduce a $0 \leqslant 0$ y es válida. Supongamos ahora $v \neq 0$. Empleando $z\bar{z} = |z|^2$ (para cualquier número complejo z) y $\langle v, u \rangle = \overline{\langle u, v \rangle}$ desarrollamos $\|u - \langle u, v \rangle tv\|^2 \geqslant 0$, donde t es cualquier número real:

$$0 \leq \| u - \langle u, v \rangle tv \|^2 = \langle u - \langle u, v \rangle tv, u - \langle u, v \rangle tv \rangle =$$
$$= \langle u, u \rangle - \overline{\langle u, v \rangle} t\langle u, v \rangle - \langle u, v \rangle t\langle v, u \rangle + \langle u, v \rangle \overline{\langle u, v \rangle} t^2 \langle v, v \rangle =$$
$$= \| u \|^2 - 2t |\langle u, v \rangle|^2 + |\langle u, v \rangle|^2 t^2 \| v \|^2$$

Tomamos $t = 1/\|v\|^2$, encontrando que $0 \leqslant \|u\|^2 - \dfrac{|\langle u, v \rangle|^2}{\|v\|^2}$, de donde $|\langle u, v \rangle|^2 \leqslant \|v\|^2 \, \|v\|^2$. Tomando la raíz cuadrada de ambos miembros obtenemos la desigualdad requerida.

6.54. Encontrar una base ortogonal para $u^\perp$ en $\mathbf{C}^3$, siendo $u = (1, i, 1 + i)$.

Aquí $u^\perp$ consiste en todos los vectores $w = (x, y, z)$ tales que

$$\langle w, u\rangle = x - iy + (1 - i)z = 0$$

Hallamos una solución, digamos $w_1 = (0, 1 - i, i)$. Después hallamos una solución del sistema

$$x - iy + (1 - i)z = 0 \qquad (1 + i)y - iz = 0$$

Aquí z es una variable libre. Tomamos $z = 1$ para obtener $y = i/(1 + i) = (1 + i)/2$ y $x = (3i - 3)2$. Multiplicando por 2 llegamos a la solución $w_2 = (3i - 3, 1 + i, 2)$. Los vectores w_1 y w_2 forman una base de $u^\perp$.

6.55. Encontrar una base ortonormal del subespacio W de $\mathbf{C}^3$ generado por

$$v_1 = (1, i, 0) \quad \text{y} \quad v_2 = (1, 2, 1 - i).$$

Aplicamos el algoritmo de Gram-Schmidt. Tomamos $w_1 = v_1 = (1, i, 0)$. Calculamos

$$v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1 = (1, 2, 1 - i) - \frac{1 - 2i}{2}(1, i, 0) = \left(\frac{1}{2} + i, 1 - \frac{1}{2}i, 1 - i\right)$$

Multiplicamos por 2 para eliminar los denominadores obteniendo $w_2 = (1 + 2i, 2 - i, 2 - 2i)$. Acto seguido hallamos $\|w_1\| = \sqrt{2}$ y entonces $\|w_2\| = \sqrt{18}$. Normalizando $\{w_1, w_2\}$ conseguimos la base ortonormal de W requerida:

$$\left\{u_1 = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right), \qquad u_2 = \left(\frac{1 + 2i}{\sqrt{18}}, \frac{2 - i}{\sqrt{18}}, \frac{2 - 2i}{\sqrt{18}}\right)\right\}$$

ESPACIOS VECTORIALES NORMADOS

6.56. Considérense los vectores $u = (1, 3, -6, 4)$ y $v = (3, -5, 1, -2)$ en $\mathbf{R}^4$. Hallar:

a) $\|u\|_\infty$ y $\|v\|_\infty$ *c)* $\|u\|_2$ y $\|v\|_2$

b) $\|u\|_1$ y $\|v\|_1$ *d)* $d_\infty(u, v)$, $d_1(u, v)$ y $d_2(u, v)$

a) La norma del supremo elige el máximo de los valores absolutos de las componentes, luego

$$\|u\|_\infty = 6 \qquad \text{y} \qquad \|v\|_\infty = 5$$

b) La 1-norma suma los valores absolutos de las componentes. Por tanto,

$$\|u\|_1 = 1 + 3 + 6 + 4 = 14 \qquad \|v\|_1 = 3 + 5 + 1 + 2 = 11$$

c) La 2-norma es igual a la raíz cuadrada de la suma de los cuadrados de las componentes (o sea, la norma inducida por el producto interno usual en $\mathbf{R}^3$). Así pues,

$$\|u\|_2 = \sqrt{1 + 9 + 36 + 16} = \sqrt{62} \qquad \text{y} \qquad \|v\|_2 = \sqrt{9 + 25 + 1 + 4} = \sqrt{39}$$

d) Primero hallamos $u - v = (-2, 8, -7, 6)$. En ese caso,

$$d_\infty(u, v) = \|u - v\|_\infty = 8$$

$$d_1(u, v) = \|u - v\|_1 = 2 + 8 + 7 + 6 = 23$$

$$d_2(u, v) = \|u - v\|_2 = \sqrt{4 + 64 + 49 + 36} = \sqrt{153}$$

6.57. Considérese la función $f(t) = t^2 - 4t$ en $C[0, 3]$. *a*) Hallar $\|f\|_\infty$. *b*) Trazar $f(t)$ en el plano $\mathbf{R}^2$. *c*) Hallar $\|f\|_1$. *d*) Hallar $\|f\|_2$.

a) Buscamos $\|f\|_\infty = \text{máx}\,(|f(t)|)$. Como $f(t)$ es diferenciable en $[0, 3]$, $|f(t)|$ alcanza un máximo en un punto crítico de $f(t)$, esto es, bien cuando la derivada $f'(t) = 0$, o bien en un extremo de $[0, 3]$. Siendo $f'(t) = 2t - 4$ hacemos $2t - 4 = 0$ y obtenemos $t = 2$ como punto crítico. Calculamos

$$f(2) = 4 - 8 = -4 \qquad f(0) = 0 - 0 = 0 \qquad f(3) = 9 - 12 = -3$$

Así $\|f\|_\infty = |f(2)| = |-4| = 4$.

b) Calculamos $f(t)$ para varios valores de t en $[0, 3]$, por ejemplo,

t	0	1	2	3
$f(t)$	0	-3	-4	-3

Marcamos los puntos en $\mathbf{R}^2$ y luego dibujamos una curva continua entre ellos, como se muestra en la Figura 6-11.

c) Buscamos $\|f\|_1 = \int_0^3 |f(t)|dt$. Como se indica en la Figura 6-11, $f(t)$ es negativa en $[0, 3]$; de aquí $|f(t)| = -(t^2 - 4t) = 4t - t^2$. De este modo,

$$\|f\|_1 = \int_0^3 (4t - t^2)\,dt = \left[2t^2 - \frac{t^3}{3}\right]_0^3 = 18 - 9 = 9$$

d)
$$\|f\|_2^2 = \int_0^3 [f(t)]^2\,dt = \int_0^3 (t^4 - 8t^3 + 16t^2)\,dt = \left[\frac{t^5}{5} - 2t^4 + \frac{16t^3}{3}\right]_0^3 = \frac{153}{5}$$

Así pues, $\|f\|_2 = \sqrt{\frac{153}{5}}$.

6.58. Demostrar el Teorema 6.25.

Si $u \neq v$, necesariamente $u - v \neq 0$, luego $d(u, v) = \|u - v\| > 0$. Además, $d(u, u) = \|u - u\| = \|0\| = 0$. Siendo así, se satisface $[M_1]$. Asimismo tenemos

$$d(u, v) = \|u - v\| = \|-1(v - u)\| = |-1|\,\|v - u\| = \|v - u\| = d(v, u)$$

y

$$d(u, v) = \|u - v\| = \|(u - w) + (w - v)\| \leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v)$$

De este modo se satisfacen $[M_2]$ y $[M_3]$.

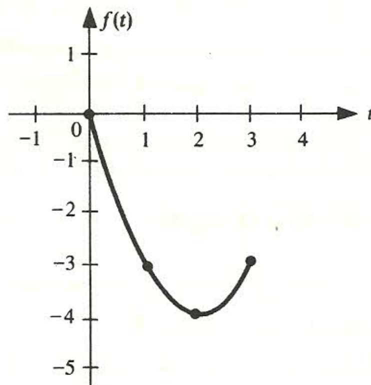


Figura 6-11.

PROBLEMAS SUPLEMENTARIOS

PRODUCTOS INTERNOS

6.59. Verificar que el siguiente es un producto interno en $\mathbf{R}^2$, donde $u = (x_1, x_2)$ y $v = (y_1, y_2)$:

$$f(u, v) = x_1 y_1 - 2x_1 y_2 - 2x_2 y_1 + 5x_2 y_2$$

6.60. Determinar los valores de k para que el que sigue sea un producto interno en $\mathbf{R}^2$, donde $u = (x_1, x_2)$ y $v = (y_1, y_2)$:

$$f(u, v) = x_1 y_1 = 3x_1 y_2 - 3x_2 y_1 + kx_2 y_2$$

6.61. Considérense los vectores $u = (1, -3)$ y $v = (2, 5)$ en $\mathbf{R}^2$. Hallar:

a) $\langle u, v\rangle$ con respecto al producto interno usual en $\mathbf{R}^2$.

b) $\langle u, v\rangle$ con respecto al producto interno en $\mathbf{R}^2$ del Problema 6.59.

c) $\|v\|$ utilizando el producto interno usual en $\mathbf{R}^2$.

d) $\|v\|$ utilizando el producto interno en $\mathbf{R}^2$ del Problema 6.59.

6.62. Probar que cada uno de los siguientes no es un producto interno en $\mathbf{R}^2$, siendo $u = (x_1, x_2, x_3)$ y $v = (y_1, y_2, y_3)$:

a) $\langle u, v\rangle = x_1 y_1 + x_2 y_2$ y *b)* $\langle u, v\rangle = x_1 y_2 x_3 + y_1 x_2 y_3$.

6.63. Sea V el espacio vectorial de las matrices $m \times n$ sobre $\mathbf{R}$. Mostrar que $\langle A, B\rangle = \operatorname{tr} B^T A$ define un producto interno en V.

6.64. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$. Probar que $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$ define un producto interno en V.

6.65. Supóngase $|\langle u, v\rangle| = \|u\| \|v\|$. (O sea, la desigualdad de Cauchy-Schwarz se reduce a una igualdad.) Probar que u y v son linealmente dependientes.

6.66. Supóngase que $f(u, v)$ y $g(u, v)$ son productos internos en un espacio vectorial V. Demostrar las siguientes aserciones:

a) La suma $f + g$ es un producto interno en V, donde $(f + g)(u, v) = f(u, v) + g(u, v)$.

b) El producto por un escalar kf, para $k > 0$, es un producto interno en V, siendo $(kf)(u, v) = kf(u, v)$.

ORTOGONALIDAD. COMPLEMENTOS ORTOGONALES. CONJUNTOS ORTOGONALES

6.67. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$ de grado $\leqslant 2$ con producto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$. Encontrar una base del subespacio W ortogonal a $h(t) = 2t + 1$.

6.68. Determinar una base del subespacio W de $\mathbf{R}^4$ ortogonal a $u_1 = (1, -2, 3, 4)$ y $u_2 = (3, -5, 7, 8)$.

6.69. Hallar una base para el subespacio W de $\mathbf{R}^5$ ortogonal a los vectores $u_1 = (1, 1, 3, 4, 1)$ y $u_2 = (1, 2, 1, 2, 1)$.

6.70. Sea $w = (1, -2, -1, 3)$ un vector en $\mathbf{R}^4$. Hallar una base: *a*) ortogonal, *b*) ortonormal de $w^\perp$.

6.71. Sea W el subespacio de $\mathbf{R}^4$ ortogonal a $u_1 = (1, 1, 2, 2)$ y $u_2 = (0, 1, 2, -1)$. Encontrar una base: *a*) ortogonal, *b*) ortonormal de W. (Compárese con el Problema 6.68.)

6.72. Supóngase que S consiste en los vectores de $\mathbf{R}^4$:

$$u_1 = (1, 1, 1, 1) \qquad u_2 = (1, 1, -1, -1) \qquad u_3 = (1, -1, 1, -1) \qquad u_4 = (1, -1, -1, 1)$$

a) Probar que S es ortogonal y es una base de $\mathbf{R}^4$.

b) Escribir $v = (1, 3, -5, 6)$ como combinación lineal de u_1, u_2, u_3, u_4.

c) Encontrar las coordenadas de un vector arbitrario $v = (a, b, c, d)$ de $\mathbf{R}^4$ relativas a la base S.

d) Normalizar S para conseguir una base ortonormal de $\mathbf{R}^4$.

6.73. Sea V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$ con producto interno $\langle A, B\rangle = \operatorname{tr}(B^T A)$. Mostrar que la que se da a continuación es una base ortonormal de V:

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}$$

6.74. Sea V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$ con producto interno $\langle A, B\rangle = \operatorname{tr} B^T A$. Hallar una base ortogonal para el complemento ortogonal de las matrices *a*) diagonales y *b*) simétricas.

6.75. Supóngase que $\{u_1, u_2, \ldots, u_r\}$ es un conjunto ortogonal de vectores. Probar que $\{k_1 u_1, k_2 u_2, \ldots, k_r u_r\}$ es ortogonal para todo conjunto de escalares $k_1, k_2, \ldots, k_r$.

6.76. Sean U y W subespacios de un espacio con producto interno V de dimensión finita. Mostrar que: *a*) $(U + W)^\perp = U^\perp \cap W^\perp$, *b*) $(U \cap W)^\perp = U^\perp + W^\perp$.

PROYECCIONES. ALGORITMO DE GRAM-SCHMIDT. APLICACIONES

6.77. Hallar una base ortogonal y una ortonormal para el subespacio U de $\mathbf{R}^4$ generado por los vectores $v_1 = (1, 1, 1, 1)$, $v_2 = (1, -1, 2, 2)$, $v_3 = (1, 2, -3, -4)$.

6.78. Sea V el espacio vectorial de los polinomios $f(t)$ con producto interno $\langle f, g\rangle = \int_0^2 f(t)g(t)dt$. Aplicar el algoritmo de Gram-Schmidt al conjunto $\{1, t, t^2\}$ para obtener un conjunto ortogonal $\{f_0, f_1, f_2\}$ con coeficientes enteros.

6.79. Supóngase $v = (1, 2, 3, 4, 6)$. Hallar la proyección de v sobre W (o encontrar un $w \in W$ que minimice $\|v - w\|$), siendo W el subespacio de $\mathbf{R}^5$ generado por:

a) $u_1 = (1, 2, 1, 2, 1)$ y $u_2 = (1, -1, 2, -1, 1)$; *b*) $v_1 = (1, 2, 1, 2, 1)$ y $v_2 = (1, 0, 1, 5, -1)$.

6.80. Sean $V = C[-1, 1]$ con producto interno $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$ y W el subespacio de V de los polinomios de grado ≤ 3. Encontrar la proyección de $f(t) = t^5$ sobre W. [*Indicación:* Utilícense los polinomios (de Legendre) $1, t, 3t^2 - 1, 5t^3 - 3t$ del Ejemplo 6.13.]

6.81. Sean $V = C[0, 1]$ con producto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$ y W el subespacio de V formado por los polinomios de grado $\leqslant 2$. Encontrar la proyección de $f(t) = t^3$ sobre W. (*Indicación:* Utilícense los polinomios 1, $2t - 1$, $6t^2 - 6t + 1$ del Problema 6.27.)

6.82. Sea U el subespacio de $\mathbf{R}^4$ generado por los vectores

$$v_1 = (1, 1, 1, 1) \qquad v_2 = (1, -1, 2, 2) \qquad v_3 = (1, 2, -3, -4)$$

Hallar la proyección de $v = (1, 2, -3, 4)$ sobre U. (*Indicación:* Hágase uso del Problema 6.77.)

PRODUCTOS INTERNOS Y MATRICES DEFINIDAS POSITIVAS. MATRICES ORTOGONALES

6.83. Determinar la matriz A que representa el producto interno usual en $\mathbf{R}^2$ respecto a cada una de las bases de $\mathbf{R}^2$: *a*) $\{v_1 = (1, 4), v_2 = (2, -3)\}$, *b*) $\{w_1 = (1, -3), w_2 = (6, 2)\}$.

6.84. Considérese el producto interno en $\mathbf{R}^2$:

$$f(u, v) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2 \qquad \text{donde} \quad u = (x_1, x_2) \quad \text{y} \quad v = (y_1, y_2)$$

Determinar la matriz B que representa este producto interno respecto a cada una de las bases del Problema 6.83.

6.85. Encontrar la matriz C que representa el producto interno usual en $\mathbf{R}^3$ respecto a la base S constituida por los vectores: $u_1 = (1, 1, 1)$, $u_2 = (1, 2, 1)$, $u_3 = (1, -1, 3)$.

6.86. Considérese el espacio vectorial V de los polinomios $f(t)$ de grado $\leqslant 2$ con producto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$.

a) Calcular $\langle f, g\rangle$, siendo $f(t) = t + 2$ y $g(t) = t^2 - 3t + 4$.

b) Hallar la matriz A del producto interno respecto a la base $(1, t, t^2)$ de V.

c) Comprobar que se verifica el Teorema 6.14, que dice que $\langle f, g\rangle = [f]^T A[g]$, respecto a la base $\{1, t, t^2\}$.

6.87. Determinar cuáles entre las matrices escritas a continuación son definidas positivas:

$$a)\ \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix}, \qquad b)\ \begin{pmatrix} 3 & 4 \\ 4 & 7 \end{pmatrix}, \qquad c)\ \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}, \qquad d)\ \begin{pmatrix} 6 & -7 \\ -7 & 9 \end{pmatrix}$$

6.88. Decidir si A es o no definida positiva, donde

$$a)\ A = \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 2 \end{pmatrix} \qquad b)\ A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 6 \\ 2 & 6 & 9 \end{pmatrix}$$

6.89. Supóngase que A y B son matrices definidas positivas. Probar que: *a*) $A + B$ es definida positiva; *b*) kA es definida positiva para $k > 0$.

6.90. Supóngase que B es una matriz real no singular. Probar que: *a*) B^TB es simétrica; *b*) B^TB es definida positiva.

6.91. Hallar el número de matrices ortogonales 2×2 de la forma $\begin{pmatrix} \frac{1}{3} & x \\ y & z \end{pmatrix}$.

6.92. Encontrar una matriz ortogonal 3×3 P cuyas dos primeras filas sean múltiplos de $u = (1, 1, 1)$ y $v = (1, -2, 3)$, respectivamente.

6.93. Encontrar una matriz ortogonal simétrica P cuya primera fila sea $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$. (Compárese con el Problema 6.43.)

6.94. Se dice que dos matrices reales A y B son *ortogonalmente equivalentes* si existe una matriz ortogonal P tal que $B = P^T AP$. Mostrar que ésta es una relación de equivalencia.

ESPACIOS COMPLEJOS CON PRODUCTO INTERNO

6.95. Comprobar que se verifica

$$\langle a_1u_1 + a_2u_2, b_1v_1 + b_2v_2\rangle = a_1\bar{b}_1\langle u_1, v_1\rangle + a_1\bar{b}_2\langle u_1, v_2\rangle + a_2\bar{b}_1\langle u_2, v_1\rangle + a_2\bar{b}_2\langle u_2, v_2\rangle$$

Con mayor generalidad, probar que $\left\langle \sum_{i=1}^{m} a_i u_i, \sum_{j=1}^{n} b_j v_j \right\rangle = \sum_{i,j} a_i \bar{b}_j \langle u_i, v_i\rangle$.

6.96. Considérense $u = (1 + i, 3, 4 - i)$ y $v = (3 - 4i, 1 + i, 2i)$ en $\mathbf{C}^3$. Calcular:

a) $\langle u, v\rangle$, *b)* $\langle v, u\rangle$, *c)* $\|u\|$, *d)* $\|v\|$, *e)* $d(u, v)$.

6.97. Hallar el coeficiente de Fourier c y la proyección cw de:

a) $u = (3 + i, 5 - 2i)$ a lo largo de $w = (5 + i, 1 + i)$ en $\mathbf{C}^2$.

b) $u = (1 - i, 3i, 1 + i)$ a lo largo de $w = (1, 2 - i, 3 + 2i)$ en $\mathbf{C}^3$.

6.98. Sean $u = (z_1, z_2)$ y $v = (w_1, w_2)$ vectores pertenecientes a $\mathbf{C}^2$. Verificar que el siguiente es un producto interno en $\mathbf{C}^2$:

$$f(u, v) = z_1\bar{w}_1 + (1 + i)z_1\bar{w}_2 + (1 - i)z_2\bar{w}_1 + 3z_2\bar{w}_2$$

6.99. Encontrar una base ortogonal y una ortonormal para el subespacio W de $\mathbf{C}^3$ generado por $u_1 = (1, i, 1)$ y $u_2 = (1 + i, 0, 2)$.

6.100. Sean $u = (z_1, z_2)$ y $v = (w_1, w_2)$ vectores pertenecientes a $\mathbf{C}^2$. ¿Para qué valores de $a, b, c, d \in \mathbf{C}$ es el que sigue un producto interno?

$$f(u, v) = az_1\bar{w}_1 + bz_1\bar{w}_2 + cz_2\bar{w}_1 + dz_2\bar{w}_2$$

6.101. Demostrar la forma polar para un espacio complejo con producto interno:

$$\langle u, v\rangle = \tfrac{1}{4}\|u + v\|^2 - \tfrac{1}{4}\|u - v\|^2 + \tfrac{1}{4}\|u + iv\|^2 - \tfrac{1}{4}\|u - iv\|^2$$

[Compárese con el Problema 6.9 *a*).]

6.102. Sea V un espacio real con producto interno. Mostrar que:

i) $\|u\| = \|v\|$ si y sólo si $\langle u + v, u - v\rangle = 0$.

ii) $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ si y sólo si $\langle u, v\rangle = 0$.

Probar, por medio de contraejemplos, que las afirmaciones anteriores no son ciertas para, por ejemplo, $\mathbf{C}^2$.

6.103. Una matriz compleja A es *unitaria* si es invertible y $A^{-1} = A^H$. Alternativamente, A es unitaria si sus filas (columnas) constituyen un conjunto ortonormal de vectores (con respecto al producto interno usual de $\mathbf{C}^n$). Encontrar una matriz unitaria cuya primera fila sea: *a*) un múltiplo de $(1, 1, -i)$; *b*) $(\frac{1}{2}, \frac{1}{2}i, \frac{1}{2} - \frac{1}{2}i)$.

ESPACIOS VECTORIALES NORMADOS

6.104. Considérense los vectores $u = (1, -3, 4, 1, -2)$ y $v = (3, 1, -2, -3, 1)$ en $\mathbf{R}^5$. Calcular:

a) $\|u\|_\infty$ y $\|v\|_\infty$ *c*) $\|u\|_2$ y $\|v\|_2$

b) $\|u\|_1$ y $\|v\|_1$ *d*) $d_\infty(u, v)$, $d_1(u, v)$ y $d_2(u, v)$

6.105. Considérense los vectores $u = (1 + i, 2 - 4i)$ y $v = (1 - i, 2 + 3i)$ en $\mathbf{C}^2$. Calcular:

a) $\|u\|_\infty$ y $\|v\|_\infty$ *c*) $\|u\|_2$ y $\|v\|_2$

b) $\|u\|_1$ y $\|v\|_1$ *d*) $d_1(u, v)$, $d_\infty(u, v)$ y $d_2(u, v)$

6.106. Considérense las funciones $f(t) = 5t - t^2$ y $g(t) = 3t - t^2$ en $C[0, 4]$. Calcular: *a*) $d_\infty(f, g)$, *b*) $d_1(f, g)$, *c*) $d_2(f, g)$.

6.107. Demostrar que: *a*) $\|\cdot\|_1$ es una norma en $\mathbf{R}^n$; *b*) $\|\cdot\|_\infty$ es una norma en $\mathbf{R}^n$.

6.108. Demostrar que: *a*) $\|\cdot\|_1$ es una norma en $C[a, b]$; *b*) $\|\cdot\|_\infty$ es una norma en $C[a, b]$.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

6.60. $k > 9$.

6.61. *a*) -13, *b*) -71, *c*) $\sqrt{29}$, *d*) $\sqrt{89}$.

6.62. Sea $u = (0, 0, 1)$. Entonces $\langle u, u \rangle = 0$ en ambos casos.

6.67. $\{7t^2 - 5t, 12t^2 - 5\}$.

6.68. $\{(1, 2, 1, 0), (4, 4, 0, 1)\}$.

6.69. $(-1, 0, 0, 0, 1), (-6, 2, 0, 1, 0), (-5, 2, 1, 0, 0)$.

6.70. *a*) $(0, 0. 3, 1), (0, 3, -3, 1), (2, 10, -9, 3)$; *b*) $(0, 0, 3, 1)/\sqrt{10}, (0, 3, -3, 1)/\sqrt{19}, (2, 10, -9, 3)/\sqrt{194}$.

6.71. *a*) $(0, 2, -1, 0), (-15, 1, 2, 5)$; *b*) $(0, 2, -1, 0)/\sqrt{5}, (-15, 1, 2, 5)/\sqrt{255}$.

6.72. *b*) $v = (5u_1 + 3u_2 - 13u_3 + 9u_4)/4$.

c) $[v] = [a + b + c + d, a + b - c - d, a - b + c - d, a - b - c + d]/4$.

6.74. *a)* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ y *b)* $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

6.77. $w_1 = (1, 1, 1, 1)$, $w_2 = (0, -2, 1, 1)$, $w_3 = (12, -4, -1, -7)$.

6.78. $f_0 = 1, f_1 = t - 1, f_3 = 3t^2 - 6t + 2$.

6.79. *a)* proy $(v, W) = (21, 27, 26, 27, 21)/8$.

b) Primero hallamos una base ortogonal de W: $w_1 = (1, 2, 1, 2, 1)$, $w_2 = (0, 2, 0, -3, 2)$. Entonces proy $(v, W) = (34, 76, 34, 56, 42)/17$.

6.80. proy $(f, W) = 10t^3/9 - 5t/21$.

6.81. proy $(f, W) = 3t^2/2 - 3t/5 + \frac{1}{20}$.

6.82. proy $(v, U) = (-14, 158, 47, 89)/70$.

6.83. *a)* $\begin{pmatrix} 17 & -10 \\ -10 & 13 \end{pmatrix}$, *b)* $\begin{pmatrix} 10 & 0 \\ 0 & 40 \end{pmatrix}$

6.84. *a)* $\begin{pmatrix} 65 & -68 \\ -68 & 73 \end{pmatrix}$, *b)* $\begin{pmatrix} 58 & 23 \\ 23 & 8 \end{pmatrix}$

6.85. $\begin{pmatrix} 3 & 4 & 3 \\ 4 & 6 & 2 \\ 3 & 2 & 11 \end{pmatrix}$

6.86. *a)* $\frac{83}{12}$, *b)* $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$

6.87. *a)* No, *b)* sí, *c)* no, *d)* sí.

6.88. *a)* Sí, *b)* no.

6.91. Cuatro: $\begin{pmatrix} \frac{1}{3} & \sqrt{8}/3 \\ \sqrt{8}/3 & -\frac{1}{3} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} & \sqrt{8}/3 \\ -\sqrt{8}/3 & -\frac{1}{3} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} & -\sqrt{8}/3 \\ \sqrt{8}/3 & \frac{1}{3} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} & -\sqrt{8}/3 \\ -\sqrt{8}/3 & -\frac{1}{3} \end{pmatrix}$

6.92. $P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{14} & -2/\sqrt{14} & 3/\sqrt{14} \\ 5/\sqrt{38} & -2/\sqrt{38} & -3/\sqrt{38} \end{pmatrix}$

6.93. $\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$

6.96. *a*) $-4i$, *b*) $4i$, *c*) $\sqrt{28}$, *d*) $\sqrt{31}$, *e*) $\sqrt{59}$.

6.97. *a*) $c = (19 - 5i)/28$, *b*) $c = (3 + 6i)/19$.

6.99. $\{v_1 = (1, i, 1)/\sqrt{3}, v_2 = (2i, 1 - 3i, 3 - i)/\sqrt{24}\}$.

6.100. a y d reales y positivos, $c = \bar{b}$ y $ad - bc$ positivo.

6.102. $u = (1, 2), v = (i, 2i)$.

6.103. *a*) $\begin{pmatrix} 1/\sqrt{3} & (1-i)/\sqrt{3} \\ (1+i)/\sqrt{3} & -1/\sqrt{3} \end{pmatrix}$, *b*) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2}i & \frac{1}{2} - \frac{1}{2}i \\ i/\sqrt{2} & -1/\sqrt{2} & 0 \\ \frac{1}{2} & -\frac{1}{2}i & -\frac{1}{2} + \frac{1}{2}i \end{pmatrix}$

6.104. *a*) 4 y 3, *b*) 11 y 13, *c*) $\sqrt{31}$ y $\sqrt{24}$, *d*) 6, 19 y 9.

6.105. *a*) $\sqrt{20}$ y $\sqrt{13}$, *b*) $\sqrt{2} + \sqrt{20}$ y $\sqrt{2} + \sqrt{13}$, *c*) $\sqrt{22}$ y $\sqrt{15}$, *d*) 7, 9 y $\sqrt{53}$.

6.106. *a*) 8, *b*) 16, *c*) $\frac{256}{3}$.

CAPITULO 7

Determinantes

7.1. INTRODUCCION

A cada matriz n-cuadrada $A = (a_{ij})$ se le asigna un escalar particular denominado el *determinante de A*, denotado por det (A), $|A|$ o

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Señalamos que una tabla ordenada $n \times n$ de escalares situada entre dos líneas verticales, llamada un *determinante de orden n*, no es una matriz. Es una forma de escribir el determinante de dicha tabla de escalares o, si se prefiere, de la matriz encerrada.

La función determinante apareció por primera vez en la investigación de los sistemas de ecuaciones lineales. Veremos que es una herramienta indispensable en el estudio y obtención de propiedades de las matrices cuadradas.

La definición de determinante y muchas de sus propiedades siguen siendo válidas cuando las entradas de la matriz proceden de un anillo.

Comenzaremos con el caso especial de los determinantes de órdenes uno, dos y tres. Posteriormente, definiremos un determinante de orden arbitrario. Esta definición general vendrá precedida por una discusión de las permutaciones, necesaria para la misma.

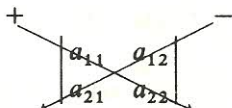
7.2. DETERMINANTES DE ORDENES UNO Y DOS

Los determinantes de órdenes uno y dos se definen como sigue:

$$|a_{11}| = a_{11}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Así el determinante de una matriz 1×1 $A = (a_{11})$ es el propio escalar a_{11}, es decir, $\det(A) = |a_{11}| = a_{11}$. El determinante de orden dos puede recordarse fácilmente usando el diagrama:



Esto es, el determinante es igual al producto de los elementos situados en la flecha marcada con un signo más menos el producto de los elementos situados en la flecha marcada con un signo menos. (Existen un diagrama análogo para determinantes de orden tres, pero no para determinantes de órdenes superiores.)

EJEMPLO 7.1

a) Dado que el determinante de orden uno es el mismo escalar, tenemos $\det(24) = 24$, $\det(-6) = -6$ y $\det(t+2) = t+2$.

b) $\begin{vmatrix} 5 & 4 \\ 2 & 3 \end{vmatrix} = (5)(3) - (4)(2) = 15 - 8 = 7$ y $\begin{vmatrix} 2 & 1 \\ -4 & 6 \end{vmatrix} = (2)(6) - (1)(-4) = 12 + 4 = 16.$

Consideremos dos ecuaciones lineales con dos incógnitas:

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

Recordemos (Problema 1.60) que el sistema tiene solución única si y sólo si $D \equiv a_1b_2 - a_2b_1 \neq 0$; y la solución es

$$x = \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1} \qquad y = \frac{a_1c_1 - a_2c_1}{a_1b_2 - a_2b_1}$$

Esta puede expresarse completamente en términos de determinantes:

$$x = \frac{N_x}{D} = \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \qquad y = \frac{N_y}{D} = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1} = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

Aquí D, el determinante de la matriz de los coeficientes, aparece en el denominador de ambos cocientes. Los numeradores N_x y N_y de los cocientes para x e y, respectivamente, pueden obtenerse sustituyendo la columna de coeficientes de la incógnita en cuestión, en la matriz de los coeficientes, por la columna de términos constantes.

EJEMPLO 7.2. Resolvamos por medio de determinantes: $\begin{cases} 2x - 3y = 7 \\ 3x + 5y = 1 \end{cases}$.

El determinante D de la matriz de los coeficientes es

$$D = \begin{vmatrix} 2 & -3 \\ 3 & 5 \end{vmatrix} = (2)(5) - (3)(-3) = 10 + 9 = 19$$

Siendo $D \neq 0$, el sistema tiene solución única. Para obtener el numerador N_x sustituyamos, en la matriz de los coeficientes, los de x por los términos constantes:

$$N_x = \begin{vmatrix} 7 & -3 \\ 1 & 5 \end{vmatrix} = (7)(5) - (1)(-3) = 35 + 3 = 38$$

Para obtener el numerador N_y sustituyamos, en la matriz de los coeficientes, los de y por los términos constantes:

$$N_y = \begin{vmatrix} 2 & 7 \\ 3 & 1 \end{vmatrix} = (2)(1) - (3)(7) = 2 - 21 = -19$$

De este modo, la solución única del sistema es

$$x = \frac{N_x}{D} = \frac{38}{19} = 2 \qquad \text{e} \qquad y = \frac{N_y}{D} = \frac{-19}{19} = -1$$

Nota: El resultado del Ejemplo 7.2 es en realidad válido para cualquier sistema de n ecuaciones con n incógnitas, como se discutirá en la Sección 7.5. Subrayamos que la importancia de este resultado es de naturaleza teórica; esto es, en la práctica, se suele utilizar la eliminación gaussiana, y no los determinantes, para hallar la solución de un sistema de ecuaciones lineales.

7.3. DETERMINANTES DE ORDEN TRES

Consideremos una matriz 3×3 arbitraria $A = (a_{ij})$. El determinante de A se define como sigue:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Obsérvese que hay seis productos, cada uno de tres elementos de la matriz original. Tres de los productos aparecen con signo más (conservan su signo), y tres con signo menos (cambian su signo).

Los diagramas de la Figura 7-1 pueden ayudar a recordar los seis productos en det (A). O sea, el determinante es igual a la suma de los productos de los elementos en las flechas marcadas con un signo más de la Figura 7-1 más la suma de los opuestos de los productos de los elementos en las flechas marcadas con un signo menos. Hacemos hincapié en que no existen tales construcciones esquemáticas para recordar los determinantes de orden superior.

EJEMPLO 7.3

$$\begin{vmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{vmatrix} = (2)(5)(4) + (1)(-2)(1) + (1)(-3)(0) - (1)(5)(1) - (-3)(-2)(2) - (4)(1)(0) =$$

$$= 40 - 2 + 0 - 5 - 12 - 0 = 21$$

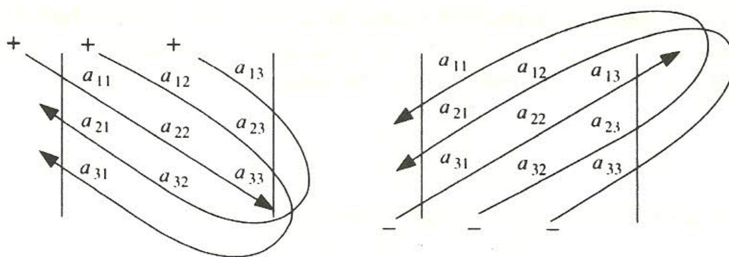


Figura 7-1.

El determinante de la matriz 3×3 $A = (a_{ij})$ puede reescribirse como:

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) =$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

que es una combinación lineal de tres determinantes de orden dos, cuyos coeficientes (con signos alternantes) constituyen la primera fila de la matriz dada. Esta combinación lineal puede indicarse de la forma

$$a_{11}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Nótese que cada matriz 2×2 se obtiene suprimiendo en la matriz inicial la fila y la columna que contienen su coeficiente.

EJEMPLO 7.4

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} = 1\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 0 & 5 & -1 \end{vmatrix} =$$

$$= 1\begin{vmatrix} -2 & 3 \\ 5 & -1 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 \\ 0 & -1 \end{vmatrix} + 3\begin{vmatrix} 4 & -2 \\ 0 & 5 \end{vmatrix} =$$

$$= 1(2 - 15) - 2(-4 + 0) + 3(20 + 0) = -13 + 8 + 60 = 55$$

7.4. PERMUTACIONES

Una permutación σ del conjunto $\{1, 2, \ldots, n\}$ es una aplicación uno-a-uno del conjunto en sí mismo o, equivalentemente, una redistribución de los números $1, 2, \ldots, n$. Tal permutación σ se denota por

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} \qquad \text{o} \qquad \sigma = j_1 j_2 \cdots j_n \qquad \text{donde } j_i = \sigma(i)$$

El conjunto de todas las permutaciones semejantes se denota por S_n, y su número es $n!$. Si $\sigma \in S_n$, la aplicación inversa $\sigma^{-1} \in S_n$; y si σ, $\tau \in S_n$, la aplicación compuesta $\sigma \circ \tau \in S_n$. Asimismo, la aplicación identidad $\varepsilon = \sigma \circ \sigma^{-1}$ pertenece a S_n. (De hecho, $\varepsilon = 12 \ldots n$.)

EJEMPLO 7.5

a) Hay $2! = 2 \cdot 1 = 2$ permutaciones en S_2: las permutaciones 12 y 21.

b) Hay $3! = 3 \cdot 2 \cdot 1 = 6$ permutaciones en S_3: las permutaciones 123, 132, 213, 231, 312, 321.

Consideremos una permutación arbitraria σ en S_n, por ejemplo, $\sigma = j_1 j_2 \ldots j_n$. Diremos que σ es una permutación par o impar según suponga un número par o impar de inversiones. Por una *inversión* en σ entendemos un par de enteros (i, k) tal que $i > k$ pero i precede a k en σ. Definimos entonces el signo o paridad de σ, escrito sgn σ, por

$$\operatorname{sgn} \sigma = \begin{cases} 1 & \text{si } \sigma \text{ es par} \\ -1 & \text{si } \sigma \text{ es impar} \end{cases}$$

EJEMPLO 7.6

a) Consideremos la permutación $\sigma = 35142$ en S_5. Para cada elemento contamos el número de elementos menores situados a su derecha. Así pues,

3 da lugar a la inversiones (3, 1) y (3, 2);
5 da lugar a las inversiones (5, 1), (5, 4), (5, 2);
4 da lugar a la invesión (4, 2).

(Adviértase que 1 y 2 no ocasionan inversiones.) Como hay, en total, seis inversiones, σ es par y sgn $\sigma = 1$.

b) La permutación identidad $\varepsilon = 12 \ldots n$ es par por carecer de inversiones.

c) En S_2 la permutación 12 es par y la 21 impar.
En S_3 las permutaciones 123, 231 y 312 son pares, mientras que las 132, 213 y 321 son impares.

d) Sea τ la permutación que intercambia dos números i y j dejando el resto de los números fijos, es decir,

$$\tau(i) = j \qquad \tau(j) = i \qquad \tau(k) = k \qquad k \neq i, j$$

Llamamos a τ trasposición. Si $i < j$,

$$\tau = 12 \ldots (i-1)j(i+1) \ldots (j-1)i(j+1) \ldots n$$

Existen $2(j - i - 1) + 1$ inversiones en τ, indicadas a continuación:

$$(j, i),\ (j, x),\ (x, i) \qquad \text{donde } x = i + 1, \ldots, j - 1$$

La trasposición τ es, pues, impar.

7.5. DETERMINANTES DE ORDEN ARBITRARIO

Sea $A = (a_{ij})$ una matriz n-cuadrada sobre un cuerpo K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Consideremos un producto de n elementos de A tal que uno y sólo uno de los elementos proviene de cada fila y uno y sólo uno proviene de cada columna. Tal producto puede escribirse:

$$a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

donde los factores proceden de filas sucesivas, de modo que los primeros subíndices mantienen el orden natural $1, 2, \dots, n$. Ahora, como los factores provienen de columnas diferentes, la sucesión de los segundos subíndices constituye una permutación $\sigma = j_1 j_2 \cdots j_n$ de S_n. Recíprocamente, cada permutación de S_n determina un producto de la forma anterior. Así pues, la matriz A origina $n!$ productos semejantes.

Definición: El determinante de $A = (a_{ij})$, denotado por $\det(A)$ o $|A|$, es la suma de los $n!$ productos precedentes, multiplicado cada uno por sgn σ. Es decir,

$$|A| = \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

o

$$|A| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

Se dice que el determinante de la matriz n-cuadrada A es de orden n.

El siguiente ejemplo muestra que esta definición concuerda con las definiciones anteriores de determinantes de órdenes uno, dos y tres.

EJEMPLO 7.7

a) Sea $A = (a_{11})$ una matriz 1×1. Puesto que S_1 sólo tiene una permutación, que es par, $\det(A) = a_{11}$, el propio número.

b) Sea $A = (a_{ij})$ una matriz 2×2. En S_2 la permutación 12 es par y la 21 impar. Por tanto,

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

c) Sea $A = (a_{ij})$ una matriz 3×3. En S_3 las permutaciones 123, 231 y 312 son pares y las 321, 213 y 132 impares. Por consiguiente,

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

A medida que n crece, el número de términos se hace astronómico. Por ello emplearemos métodos indirectos para evaluar los determinantes más que su definición. De hecho demostraremos una serie de propiedades de los determinantes que nos permitirán abreviar los cálculos considerablemente. En particular probaremos que un determinante de orden n es igual a una combinación lineal de determinantes de orden $n - 1$, como ocurría en el caso $n = 3$.

7.6. PROPIEDADES DE LOS DETERMINANTES

Enumeramos a continuación las propiedades básicas del determinante.

Teorema 7.1: El determinante de una matriz A y el de su traspuesta A^T son iguales; o sea, $|A| = |A^T|$.

De acuerdo con este resultado, demostrado en el Problema 7.21, cualquier teorema relativo a una matriz A que concierna a sus filas tendrá un análogo concerniente a sus columnas.

El próximo teorema, demostrado en el Problema 7.23, nos descubre ciertas situaciones en las que el determinante puede obtenerse de manera inmediata.

Teorema 7.2: Sea A una matriz cuadrada.

a) Si A posee una fila (columna) de ceros, necesariamente $|A| = 0$.

b) Si A posee dos filas (columnas) iguales, necesariamente $|A| = 0$.

c) Si A es triangular, esto es, A tiene sólo ceros por encima o por debajo de la diagonal, entonces $|A|$ es igual al producto de los elementos diagonales. De este modo, en particular, $|I| = 1$, siendo I la matriz identidad.

El siguiente teorema, demostrado en el Problema 7.22, muestra cómo se ve afectado el determinante de una matriz por las operaciones elementales entre filas o entre columnas.

Teorema 7.3: Supongamos que B se ha obtenido de A mediante una operación elemental entre filas (columnas).

a) Si se han intercambiado dos filas (columnas) de A, $|B| = -|A|$.

b) Si se ha multiplicado una fila (columna) de A por un escalar k, $|B| = k|A|$.

c) Si se ha sumado un múltiplo de una fila (columna) a otra, $|B| = |A|$.

Ahora enunciamos dos de los teoremas más importantes y útiles referentes a determinantes.

Teorema 7.4: Sea A cualquier matriz n-cuadrada. Son equivalentes las siguientes aserciones:

i) A es invertible, es decir, A tiene una inversa A^{-1}.

ii) $AX = 0$ tiene únicamente la solución trivial.

iii) El determinante de A es no nulo: $|A| \neq 0$.

Nota: Dependiendo del autor y del texto, una matriz no singular A se define como una matriz invertible, como una para la que $|A| \neq 0$, o como una para la que $AX = 0$ sólo tiene la solución nula. El teorema anterior muestra que las tres definiciones son equivalentes.

Teorema 7.5: El determinante es una función multiplicativa. O sea, el determinante del producto de dos matrices A y B es el producto de los determinantes: $|AB| = |A||B|$.

Demostraremos los dos teoremas precedentes (Problemas 7.27 y 7.28, respectivamente) utilizando la teoría de las matrices elementales, junto al lema que sigue (demostrado a su vez en el Problema 7.25).

Lema 7.6: Sea E una matriz elemental. En tal caso, para toda matriz A, $|EA| = |E||A|$.

Cabe comentar que es posible probar los dos teoremas directamente, sin necesidad de recurrir a la teoría de las matrices elementales.

Recordemos que dos matrices A y B son similares si existe una matriz no singular P tal que $B = P^{-1}AP$. Haciendo uso de la propiedad multiplicativa del determinante (Teorema 7.5) seremos capaces de demostrar (Problema 7.30):

Teorema 7.7: Supongamos que A y B son matrices similares. Entonces $|A| = |B|$.

7.7. MENORES Y COFACTORES

Consideremos una matriz n-cuadrada $A = (a_{ij})$. Denotemos por M_{ij} la submatriz $(n-1)$-cuadrada de A que se obtiene suprimiendo su i-ésima fila y su j-ésima columna. El determinante $|M_{ij}|$ recibe el nombre de menor del elemento a_{ij} de A, y definimos el cofactor de a_{ij}, denotado por A_{ij}, como el menor «con signo»:

$$A_{ij} = (-1)^{i+j}|M_{ij}|$$

Nótese que los «signos» $(-1)^{i+j}$ que acompañan a los menores se disponen en forma de tablero de ajedrez, con signos $+$ en la diagonal principal:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Subrayamos el hecho de que M_{ij} denota una matriz, mientras que A_{ij} denota un escalar.

Nota: El signo $(-1)^{i+j}$ del cofactor A_{ij} se obtiene frecuentemente utilizando la configuración de tablero de ajedrez citada. Concretamente, empezando con «$+$» y alternando los signos, esto es, $+, -, +, -, \ldots$, se cuenta desde la diagonal principal hasta la casilla apropiada.

EJEMPLO 7.8. Consideremos la matriz $A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix}$.

$$|M_{23}| = \begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} = 18 - 24 = -6 \text{ y por tanto } A_{23} = (-1)^{2+3}|M_{23}| = (-1)\cdot(-6) = 6.$$

Es aplicable el teorema enunciado a continuación, demostrado en el Problema 7.31.

Teorema 7.8. El determinante de la matriz $A = (a_{ij})$ es igual a la suma de los productos obtenidos multiplicando los elementos de cualquier fila (columna) por sus correspondientes cofactores:

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^{n} a_{ij}A_{ij}$$

y

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}$$

Estas fórmulas para $|A|$ se conocen como *desarrollos de Laplace* del determinante de A por la i-ésima fila y la j-ésima columna, respectivamente. Junto con las operaciones elementales entre filas, ofrecen un método de simplificación del cálculo de $|A|$, como se describe más abajo.

EVALUACION DE DETERMINANTES

El siguiente algoritmo reduce la evaluación de un determinante de orden n a la de uno de orden $n - 1$.

Algoritmo 7.7 (Reducción del orden de un determinante)

Aquí $A = (a_{ij})$ es una matriz n-cuadrada no nula, con $n > 1$.

Paso 1. Elegir un elemento $a_{ij} = 1$ o, en su defecto, un $a_{ij} \neq 0$.

Paso 2. Usando a_{ij} como pivote, efectuar las operaciones elementales entre filas [columnas] necesarias para colocar ceros en el resto de las posiciones de la columna [fila] que contiene a_{ij}.

Paso 3. Desarrollar el determinante por la columna [fila] que contiene a_{ij}.

Podemos hacer, por orden, las siguientes observaciones.

Nota 1: El Algoritmo 7.7 suele emplearse para determinantes de orden cuatro o mayor. Con determinantes de orden menor que cuatro se utilizan las fórmulas del determinante específicas.

Nota 2: La eliminación gaussiana o, equivalentemente, el uso repetido del Algoritmo 7.7 acompañado de intercambios de filas puede servir para transformar una matriz A en otra triangular superior, cuyo determinante es el producto de las entradas diagonales. No obstante, no debe perderse de vista el número de intercambios de filas, pues cada uno de ellos varía el signo del determinante. (Véase el Problema 7.11.)

EJEMPLO 7.9. Calculemos el determinante de $A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}$ por medio del Algoritmo 7.7.

Utilizamos $a_{23} = 1$ como pivote para situar ceros en el resto de las posiciones de la tercera columna, esto es, efectuamos las operaciones entre filas $-2R_2 + R_1 \to R_1$, $3R_2 + R_3 \to R_3$ y $R_2 + R_4 \to R_4$. Según el Teorema 7.3 *c*), el valor del determinante no se altera con estas operaciones, es decir,

$$|A| = \begin{vmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix}$$

Si ahora desarrollamos por la tercera columna, podemos despreciar todos los términos que contienen 0. Así

$$|A| = (-1)^{2+3} \begin{vmatrix} 1 & -2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & -2 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix} = -(4 - 18 + 5 - 30 - 3 + 4) = -(-38) = 38$$

7.8. ADJUNTO CLASICO

Consideremos una matriz n-cuadrada $A = (a_{ij})$ sobre un cuerpo K. El adjunto clásico (tradicionalmente llamado «adjunto») de A, denotado por adj A, es la traspuesta de la matriz de cofactores de A:

$$\text{adj } A = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

Decimos «adjunto clásico» en vez de simplemente «adjunto» debido a que este último término se suele usar para un concepto completamente diferente.

EJEMPLO 7.10. Sea $A = \begin{pmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{pmatrix}$. Los cofactores de los nueve elementos de A son

$$A_{11} = +\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -18 \qquad A_{12} = -\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} = 2 \qquad A_{13} = +\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} = 4$$

$$A_{21} = -\begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} = -11 \qquad A_{22} = +\begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} = 14 \qquad A_{23} = -\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 5$$

$$A_{31} = +\begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} = -10 \qquad A_{32} = -\begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} = -4 \qquad A_{33} = +\begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8$$

La traspuesta de la matriz de los cofactores anteriores proporciona el adjunto clásico de A:

$$\text{adj } A = \begin{pmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{pmatrix}$$

Es aplicable el siguiente teorema, demostrado en el Problema 7.33.

Teorema 7.9: Para toda matriz cuadrada A,

$$A \cdot (\text{adj } A) = (\text{adj } A) \cdot A = |A| I$$

donde I es la matriz identidad. De este modo, si $|A| \neq 0$,

$$A^{-1} = \frac{1}{|A|} (\text{adj } A)$$

Observemos que este teorema pone en nuestras manos otro método para hallar la inversa de una matriz no singular dada.

EJEMPLO 7.11. Consideremos la matriz A del Ejemplo 7.10 para la cual $|A| = -46$. Tenemos

$$A \cdot (\text{adj } A) = \begin{pmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{pmatrix} = \begin{pmatrix} -46 & 0 & 0 \\ 0 & -46 & 0 \\ 0 & 0 & -46 \end{pmatrix} = -46 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$= -46 I = |A| I$$

Además, de acuerdo con el Teorema 7.9,

$$A^{-1} = \frac{1}{|A|} (\text{adj } A) = \begin{pmatrix} -18/(-46) & -11/(-46) & -10/(-46) \\ 2/(-46) & 14/(-46) & -4/(-46) \\ 4/(-46) & 5/(-46) & -8/(-46) \end{pmatrix} = \begin{pmatrix} \frac{9}{23} & \frac{11}{46} & \frac{5}{23} \\ -\frac{1}{23} & -\frac{7}{23} & \frac{2}{23} \\ -\frac{2}{23} & -\frac{5}{46} & \frac{4}{23} \end{pmatrix}$$

7.9. APLICACIONES A LAS ECUACIONES LINEALES. REGLA DE CRAMER

Consideremos un sistema de n ecuaciones lineales con n incógnitas:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}$$

El sistema puede escribirse en la forma $AX = B$, donde $A = (a_{ij})$ es la matriz (cuadrada) de coeficientes y $B = (b_i)$ el vector columna de constantes. Sea A_i la matriz obtenida a partir de A sustituyendo su columna i-ésima por B. Sean $D = \det(A)$ y $N_i = \det(A_i)$ para $i = 1, 2, \ldots, n$. La relación fundamental entre los determinantes y la solución del sistema es la siguiente.

Teorema 7.10: El sistema precedente tiene solución única si y sólo si $D \neq 0$, en cuyo caso la solución viene dada por

$$x_1 = N_1/D, \qquad x_2 = N_2/D, \qquad \ldots, \qquad x_n = N_n/D$$

Este teorema, demostrado en el Problema 7.34, se conoce como «regla de Cramer» para la resolución de sistemas de ecuaciones lineales. Hacemos énfasis en que el teorema sólo se refiere a sistemas con el mismo número de ecuaciones que de incógnitas, y que únicamente da la solución cuando $D \neq 0$. De hecho, si $D = 0$, el teorema no nos dice si el sistema es o no compatible. A pesar de ello, en el caso de un sistema homogéneo, disponemos del resultado, de gran utilidad, que se escribe a continuación (y se demuestra en el Problema 7.65).

Teorema 7.11: El sistema homogéneo $AX = 0$ tiene solución no nula si y sólo si $D = |A| = 0$.

EJEMPLO 7.12. Resolvemos, empleando determinantes: $\begin{cases} 2x + y - z = 3 \\ x + y + z = 1. \\ x - 2y - 3z = 4 \end{cases}$

Comenzamos calculando el determinante D de la matriz de los coeficientes:

$$D = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -2 & -3 \end{vmatrix} = -6 + 1 + 2 + 1 + 4 + 3 = 5$$

Como $D \neq 0$, el sistema tiene solución única. Para encontrar N_x, N_y y N_z sustituimos los coeficientes de x, y y z en la matriz por los términos constantes:

$$N_x = \begin{vmatrix} 3 & 1 & -1 \\ 1 & 1 & 1 \\ 4 & -2 & -3 \end{vmatrix} = -9 + 4 + 2 + 4 + 6 + 3 = 10$$

$$N_y = \begin{vmatrix} 2 & 3 & -1 \\ 1 & 1 & 1 \\ 1 & 4 & -3 \end{vmatrix} = -6 + 3 - 4 + 1 - 8 + 9 = -5$$

$$N_z = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 1 & 1 \\ 1 & -2 & 4 \end{vmatrix} = 8 + 1 - 6 - 3 + 4 - 4 = 0$$

La solución única es, pues, $x = N_x/D = 2$, $y = N_y/D = -1$, $z = N_z/D = 0$.

7.10. SUBMATRICES. MENORES GENERALES. MENORES PRINCIPALES

Sea $A = (a_{ij})$ una matriz n-cuadrada. Cada par de conjuntos ordenados $i_1, i_2, \ldots, i_r$ de r índices de fila y $j_1, j_2, \ldots, j_r$ de r índices de columna define la submatriz de A de orden r:

$$A_{i_1, i_2, \ldots, i_r}^{j_1, j_2, \ldots, j_r} = \begin{pmatrix} a_{i_1, j_1} & a_{i_1, j_2} & \cdots & a_{i_1, j_r} \\ a_{i_2, j_1} & a_{i_2, j_2} & \cdots & a_{i_2, j_2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_r, j_1} & a_{i_r, j_2} & \cdots & a_{i_r, j_r} \end{pmatrix}$$

El determinante $|A_{i_1, i_2, \ldots, i_r}^{j_1, j_2, \ldots, j_r}|$ se denomina un menor de A de orden r, y

$$(-1)^{i_1 + i_2 + \cdots + i_r + j_1 + j_2 + \cdots + j_r} |A_{i_1, i_2, \ldots, i_r}^{j_1, j_2, \ldots, j_r}|$$

es el menor con signo correspondiente. (Nótese que un menor de orden $n - 1$ es un menor en el sentido de la Sección 7.7, siendo el correspondiente menor con signo un cofactor.) Además, si i'_k y j'_k denotan, respectivamente, los índices de fila y columna restantes,

$$|A_{i'_1, \ldots, i'_{n-r}}^{j'_1, \ldots, j'_{n-r}}|$$

es el *menor complementario*.

EJEMPLO 7.13. Supongamos que $A = (a_{ij})$ es una matriz 5-cuadrada. Los subíndices de fila 3 y 5 y los de columna 1 y 4 definen la submatriz

$$A_{3, 5}^{1, 4} = \begin{pmatrix} a_{31} & a_{34} \\ a_{51} & a_{54} \end{pmatrix}$$

y el menor y menor con signo son, respectivamente,

$$|A_{3, 5}^{1, 4}| = \begin{vmatrix} a_{31} & a_{34} \\ a_{51} & a_{54} \end{vmatrix} = a_{31}a_{54} - a_{34}a_{51} \qquad \text{y} \qquad (-1)^{3+5+1+4}|A_{3, 5}^{1, 4}| = -|A_{3, 5}^{1, 4}|$$

Los subíndices de fila restantes son 1, 2 y 4 y los de columna 2, 3 y 5, luego el menor complementario de $|A_{3, 5}^{1, 4}|$ es

$$|A_{1, 2, 4}^{2, 3, 5}| = \begin{vmatrix} a_{12} & a_{13} & a_{15} \\ a_{22} & a_{23} & a_{35} \\ a_{42} & a_{43} & a_{45} \end{vmatrix}$$

MENORES PRINCIPALES

Un menor es principal si los índices de fila y de columna coinciden o, en otras palabras, si los elementos diagonales del menor proceden de la diagonal de la matriz.

EJEMPLO 7.14. Consideremos los menores de una matriz 5-cuadrada $A = (a_{ij})$:

$$M_1 = \begin{vmatrix} a_{22} & a_{24} & a_{25} \\ a_{42} & a_{44} & a_{45} \\ a_{52} & a_{54} & a_{55} \end{vmatrix} \qquad M_2 = \begin{vmatrix} a_{11} & a_{13} & a_{15} \\ a_{21} & a_{23} & a_{25} \\ a_{51} & a_{53} & a_{55} \end{vmatrix} \qquad M_3 = \begin{vmatrix} a_{22} & a_{25} \\ a_{52} & a_{55} \end{vmatrix}$$

Aquí M_1 y M_3 son menores principales, porque todos sus elementos diagonales pertenecen a la diagonal de A. Sin embargo, M_2 no es principal, ya que a_{23} pertenece a la diagonal de M_2 pero no a la de A.

Se pueden hacer, por orden, las siguientes observaciones.

Nota 1: El signo de un menor principal es siempre $+1$, puesto que la suma de los índices de fila y los idénticos de columna es par.

Nota 2: Un menor es principal si y sólo si su menor complementario es también principal.

Nota 3: Una matriz real simétrica A es definida positiva si y sólo si todos sus menores principales son positivos.

7.11. MATRICES POR BLOQUES Y DETERMINANTES

A continuación se expone el resultado fundamental de esta sección.

Teorema 7.12: Supongamos que M es una matriz triangular superior (inferior) por bloques, con bloques diagonales $A_1, A_2, \ldots, A_n$. Entonces

$$\det(M) = \det(A_1)\det(A_2)\cdots\det(A_n)$$

La demostración aparece en el Problema 7.35.

EJEMPLO 7.15. Hallemos $|M|$, siendo $M = \left(\begin{array}{cc|ccc} 2 & 3 & 4 & 7 & 8 \\ -1 & 5 & 3 & 2 & 1 \\ \hline 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 3 & -1 & 4 \\ 0 & 0 & 5 & 2 & 6 \end{array}\right)$.

Nótese que M es una matriz triangular superior por bloques. Evaluamos el determinante de cada bloque diagonal:

$$\begin{vmatrix} 2 & 3 \\ -1 & 5 \end{vmatrix} = 10 + 3 = 13 \qquad \begin{vmatrix} 2 & 1 & 5 \\ 3 & -1 & 4 \\ 5 & 2 & 6 \end{vmatrix} = -12 + 20 + 30 + 25 - 16 - 18 = 29$$

Por tanto, $|M| = (13)(29) = 377$.

Nota: Supongamos $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, donde A, B, C, D son matrices cuadradas. En general no será cierto $|M| = |A|\,|D| - |B|\,|C|$. (Véase el Problema 7.77.)

7.12. DETERMINANTES Y VOLUMEN

Los determinantes están ligados a las nociones de área y volumen como sigue. Sean $u_1, u_2, \ldots, u_n$ vectores en $\mathbf{R}^n$ y S el paralelepípedo que determinan; esto es,

$$S = \{a_1u_1 + a_2u_2 + \cdots + a_nu_n : 0 \le a_i \le 1 \text{ para } i = 1, \ldots, n\}$$

(Cuando $n = 2$, S es un paralelogramo.) Denotemos por $V(S)$ el volumen (o área, si $n = 2$) de S. En tal caso,

$$V(S) = \text{valor absoluto de det } (A)$$

donde A es la matriz con filas $u_1, u_2, \ldots, u_n$. En general, $V(S) = 0$ si y sólo si los vectores $u_1, u_2, \ldots, u_n$ no constituyen un sistema de coordenadas de $\mathbf{R}^n$, es decir, si y sólo si los vectores son linealmente dependientes.

7.13. MULTILINEALIDAD Y DETERMINANTES

Sean V un espacio vectorial sobre un cuerpo K y $\mathscr{A} = V^n$, esto es, $\mathscr{A}$ consiste en todas las n-plas

$$A = (A_1, A_2, \ldots, A_n)$$

en las que los A_i son vectores de V. Son pertinentes las definiciones:

Definición: Se dice que una función $D: \mathscr{A} \to K$ es *multilineal* si es lineal en cada uno de sus argumentos; o sea:

i) Si $A_i = B + C$,

$$D(A) = D(\ldots, B + C, \ldots) = D(\ldots, B, \ldots) + D(\ldots, C, \ldots)$$

ii) Si $A_i = kB$ con $k \in K$,

$$D(A) = D(\ldots, kB, \ldots) = kD(\ldots, B, \ldots)$$

También se utiliza la expresión n-lineal en lugar de multilineal cuando hay n argumentos.

Definición: Se dice que una función $D: \mathscr{A} \to K$ es alternada si $D(A) = 0$ siempre que A tenga dos elementos iguales; es decir,

$$D(A_1, A_2, \ldots, A_n) = 0 \qquad \text{siempre que} \qquad A_i = A_j,\ i \neq j$$

Denotemos ahora por $\mathbf{M}$ el conjunto de todas las matrices n-cuadradas A sobre un cuerpo K. Podemos ver A como una n-pla constituida por sus vectores fila $A_1, A_2, \ldots, A_n$; esto es, podemos ver A de la forma $A = (A_1, A_2, \ldots, A_n)$.

El siguiente resultado básico (demostrado en el Problema 7.36) es aplicable. (I denota la matriz identidad.)

Teorema 7.13: Existe una única función $D: \mathbf{M} \to K$ tal que

i) D es multilineal, ii) D es alternada, iii) $D(I) = 1$

Esta función D no es otra que la función determinante; es decir, para toda matriz $A \in \mathbf{M}$, $D(A) = |A|$.

PROBLEMAS RESUELTOS

CALCULO DE DETERMINANTES DE ORDENES DOS Y TRES

7.1. Evaluar el determinante de cada una de las matrices:

$$\begin{pmatrix} 6 & 5 \\ 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 3 & -2 \\ 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 4 & -5 \\ -1 & -2 \end{pmatrix}$$

$$\begin{vmatrix} 6 & 5 \\ 2 & 3 \end{vmatrix} = (6)(3) - (5)(2) = 18 - 10 = 8, \qquad \begin{vmatrix} 3 & -2 \\ 4 & 5 \end{vmatrix} = 15 + 8 = 23, \qquad \begin{vmatrix} 4 & -5 \\ -1 & -2 \end{vmatrix} = -8 - 5 = -13$$

7.2. Hallar el determinante de $\begin{pmatrix} t-5 & 7 \\ -1 & t+3 \end{pmatrix}$.

$$\begin{vmatrix} t-5 & 7 \\ -1 & t+3 \end{vmatrix} = (t-5)(t+3) + 7 = t^2 - 2t - 15 + 7 = t^2 - 2t - 8$$

7.3. Determinar los valores de k para los que $\begin{vmatrix} k & k \\ 4 & 2k \end{vmatrix} = 0$.

Hacemos $\begin{vmatrix} k & k \\ 4 & 2k \end{vmatrix} = 2k^2 - 4k = 0$ o $2k(k-2) = 0$. De aquí $k = 0$ y $k = 2$. O sea, si $k = 0$ o $k = 2$, el determinante es nulo.

7.4. Evaluar los determinantes de las siguientes matrices:

a) $\begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \\ 1 & 5 & -2 \end{pmatrix}$, *b)* $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$

Utilícense los diagramas de la Figura 7-1.

a) $$\begin{vmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \\ 1 & 5 & -2 \end{vmatrix} = (1)(4)(-2) + (-2)(-1)(1) + (3)(5)(2) - (1)(4)(3) - (5)(-1)(1) - (-2)(-2)(2) = -8 + 2 + 30 - 12 + 5 - 8 = 9$$

b) $$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 b_3 a_2 - a_3 b_2 c_1 - b_3 c_2 a_1 - c_3 b_1 a_2$$

7.5. Calcular el determinante de $\begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix}$.

Comenzamos por simplificar las entradas restando dos veces la primera fila a la segunda, esto es, efectuando $-2R_1 + R_2 \to R_2$:

$$\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 \\ 1 & 0 & -1 \\ 8 & 9 & 1 \end{vmatrix} = 0 - 24 + 36 - 0 + 18 - 3 = 27$$

7.6. Hallar el determinante de A, donde:

a) $A = \begin{pmatrix} \frac{1}{2} & -1 & -\frac{1}{3} \\ \frac{3}{4} & \frac{1}{2} & -1 \\ 1 & -4 & 1 \end{pmatrix}$, *b*) $A = \begin{pmatrix} t+3 & -1 & 1 \\ 5 & t-3 & 1 \\ 6 & -6 & t+4 \end{pmatrix}$

a) Primero multiplicamos la primera fila por 6 y la segunda por 4. Entonces

$$6 \cdot 4|A| = 24|A| = \begin{vmatrix} 3 & -6 & -2 \\ 3 & 2 & -4 \\ 1 & -4 & 1 \end{vmatrix} = 6 + 24 + 24 + 4 - 48 + 18 = 28$$

Por tanto, $|A| = \frac{28}{24} = \frac{7}{6}$. (Obsérvese que las multiplicaciones iniciales eliminaron los denominadores, simplificando la aritmética.)

b) Sumamos la segunda columna a la primera, y después la tercera a la segunda para conseguir ceros; es decir, efectuamos $C_2 + C_1 \to C_1$ y $C_3 + C_2 \to C_2$:

$$|A| = \begin{vmatrix} t+2 & 0 & 1 \\ t+2 & t-2 & 1 \\ 0 & t-2 & t+4 \end{vmatrix}$$

Ahora sacamos el factor $t + 2$ de la primera columna y el $t - 2$ de la segunda para llegar a

$$|A| = (t+2)(t-2)\begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & t+4 \end{vmatrix}$$

Finalmente, restamos la primera columna a la tercera obteniendo

$$|A| = (t+2)(t-2)\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & t+4 \end{vmatrix} = (t+2)(t-2)(t+4)$$

CALCULO DE DETERMINANTES DE ORDENES ARBITRARIOS

7.7. Calcular el determinante de $A = \begin{pmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{pmatrix}$.

Usamos $a_{31} = 1$ como pivote y efectuamos las operaciones entre filas $-2R_3 + R_1 \to R_1$, $2R_3 + R_1 \to R_2$ y $R_3 + R_4 \to R_4$:

$$|A| = \begin{vmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 1 & -6 \\ 0 & 3 & -2 & -1 \\ 1 & 3 & -2 & 2 \\ 0 & -3 & 2 & 5 \end{vmatrix} = + \begin{vmatrix} -1 & 1 & -6 \\ 3 & -2 & -1 \\ -3 & 2 & 5 \end{vmatrix} =$$

$$= 10 + 3 - 36 + 36 - 2 - 15 = -4$$

7.8. Evaluar el determinante de $A = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 3 & -1 & 1 & -2 \\ 4 & -3 & 0 & 2 \end{pmatrix}$.

Usamos $a_{21} = 1$ como pivote y efectuamos $2C_1 + C_3 \to C_3$:

$$|A| = \begin{vmatrix} 1 & 2 & 4 & 3 \\ 1 & 0 & 0 & 0 \\ 3 & -1 & 7 & -2 \\ 4 & -3 & 6 & 2 \end{vmatrix} = - \begin{vmatrix} 2 & 4 & 3 \\ -1 & 7 & -2 \\ -3 & 8 & 2 \end{vmatrix} = -(28 + 24 - 24 + 63 + 32 + 8) = -131$$

7.9. Hallar el determinante de $C = \begin{pmatrix} 6 & 2 & 1 & 0 & 5 \\ 2 & 1 & 1 & -2 & 1 \\ 1 & 1 & 2 & -2 & 3 \\ 3 & 0 & 2 & 3 & -1 \\ -1 & -1 & -3 & 4 & 2 \end{pmatrix}$.

Empezamos reduciendo $|C|$ a un determinante de orden cuatro y luego a uno de orden tres. Empleamos $c_{22} = 1$ como pivote y efectuamos $-2R_2 + R_1 \to R_1$, $-R_2 + R_3 \to R_3$ y $R_2 + R_5 \to R_5$:

$$|C| = \begin{vmatrix} 2 & 0 & -1 & 4 & 3 \\ 2 & 1 & 1 & -2 & 1 \\ -1 & 0 & 1 & 0 & 2 \\ 3 & 0 & 2 & 3 & -1 \\ 1 & 0 & -2 & 2 & 6 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 4 & 3 \\ -1 & 1 & 0 & 2 \\ 3 & 2 & 3 & -1 \\ 1 & -2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 4 & -1 \\ 0 & 1 & 0 & 0 \\ 5 & 2 & 3 & -5 \\ -1 & -2 & 2 & 7 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 4 & -1 \\ 5 & 3 & -5 \\ -1 & 2 & 7 \end{vmatrix} = 21 + 20 - 10 - 3 + 10 - 140 = -102$$

7.10. Hallar el determinante de cada una de las siguientes matrices:

$$a)\quad A = \begin{pmatrix} 5 & 6 & 7 & 8 \\ 0 & 0 & 0 & 0 \\ 1 & -3 & 5 & -7 \\ 8 & 4 & 2 & 6 \end{pmatrix}, \qquad b)\quad B = \begin{pmatrix} 5 & 6 & 7 & 6 \\ 1 & -3 & 5 & -3 \\ 4 & 9 & -3 & 9 \\ 2 & 7 & 8 & 7 \end{pmatrix},$$

$$c)\quad C = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & -3 & 7 & -8 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

a) Como A tiene una fila de ceros, $\det(A) = 0$.

b) Siendo iguales las columnas segunda y cuarta de B, $\det(B) = 0$.

c) Dado que C es triangular, $\det(C)$ es igual al producto de las entradas diagonales. Por consiguiente, $\det(C) = -120$.

7.11. Describir el algoritmo de eliminación gaussiana para el cálculo del determinante de una matriz n-cuadrada $A = (a_{ij})$.

El algoritmo utiliza la eliminación gaussiana para transformar A en una matriz triangular superior (cuyo determinante es el producto de las entradas diagonales). Como el algoritmo implica intercambios de filas, con los consiguientes cambios de signo del determinante, debe seguirse la pista de tales cambios empleando para ello alguna variable, digamos SGN. El algoritmo también hace uso del mecanismo de «pivotar»; esto es, el elemento de mayor valor absoluto será el que se utilice como pivote. Exponemos el algoritmo a continuación.

Paso 1. Tomar SGN = 0 [Con esto se inicializa la variable SGN.]

Paso 2. Hallar la entrada de la primera columna a_{i1} de mayor valor absoluto.

a) Si $a_{i1} = 0$, hacer $\det(A) = 0$ y SALIR.

b) Si $i \neq 1$, intercambiar las filas primea e i-ésima y hacer SGN = SGN + 1.

Paso 3. Usar a_{11} como pivote y efectuar operaciones elementales entre filas de la forma $kR_q + R_p \to R_p$ para situar ceros bajo a_{11}.

Paso 4. Repetir los Pasos 2 y 3 con la submatriz obtenida omitiendo la primera fila y columna.

Paso 5. Continuar el proceso anterior hasta que A sea una matriz triangular superior.

Paso 6. Hacer $\det(A) = (-1)^{\text{SGN}} a_{11} a_{22} \cdots a_{nn}$ y SALIR.

Nótese que la operación elemental entre filas de la forma $kR_p \to R_p$ (que multiplica una fila por un escalar), permitida en el algoritmo de eliminación gaussiana para un sistema de ecuaciones lineales, está prohibida aquí, ya que modifica el valor del determinante.

7.12. Utilizando el algoritmo de eliminación gaussiana del Problema 7.11, hallar el determinante de

$$A = \begin{pmatrix} 3 & 8 & 6 \\ -2 & -3 & 1 \\ 5 & 10 & 15 \end{pmatrix}$$

Primero reducimos la matriz A a una triangular superior sin perder de vista el número de intercambios de filas:

$$A \sim \begin{pmatrix} 5 & 10 & 15 \\ -2 & -3 & 1 \\ 3 & 8 & 6 \end{pmatrix} \sim \begin{pmatrix} 5 & 10 & 15 \\ 0 & 1 & 7 \\ 0 & 2 & -3 \end{pmatrix} \sim \begin{pmatrix} 5 & 10 & 15 \\ 0 & 2 & -3 \\ 0 & 1 & 7 \end{pmatrix} \sim \begin{pmatrix} 5 & 10 & 15 \\ 0 & 2 & -3 \\ 0 & 0 & \frac{17}{2} \end{pmatrix}$$

Ahora A está forma triangular y SGN = 2 porque ha habido dos intercambios de filas. Por tanto, $A = (-1)^{\text{SGN}} \cdot 5 \cdot 2 \cdot (\frac{17}{2}) = 85$.

7.13. Evaluar $|B| = \begin{vmatrix} 0.921 & 0.185 & 0.476 & 0.614 \\ 0.782 & 0.157 & 0.527 & 0.138 \\ 0.872 & 0.484 & 0.637 & 0.799 \\ 0.312 & 0.555 & 0.841 & 0.448 \end{vmatrix}$.

Multiplicamos la fila que contiene el pivote a_{ij} por $1/a_{ij}$, de forma que el pivote sea igual a 1:

$$|B| = 0.921 \begin{vmatrix} 1 & 0.201 & 0.517 & 0.667 \\ 0.782 & 0.157 & 0.527 & 0.138 \\ 0.872 & 0.484 & 0.637 & 0.799 \\ 0.312 & 0.555 & 0.841 & 0.448 \end{vmatrix} = 0.921 \begin{vmatrix} 1 & 0.201 & 0.571 & 0.667 \\ 0 & 0 & 0.123 & -0.384 \\ 0 & 0.309 & 0.196 & 0.217 \\ 0 & 0.492 & 0.680 & 0.240 \end{vmatrix} =$$

$$= 0.921 \begin{vmatrix} 0 & 0.123 & -0.384 \\ 0.309 & 0.196 & 0.217 \\ 0.492 & 0.680 & 0.240 \end{vmatrix} = 0.921(-0.384) \begin{vmatrix} 0 & -0.320 & 1 \\ 0.309 & 0.196 & 0.217 \\ 0.492 & 0.680 & 0.240 \end{vmatrix} =$$

$$= 0.921(-0.384) \begin{vmatrix} 0 & 0 & 1 \\ 0.309 & 0.265 & 0.217 \\ 0.492 & 0.757 & 0.240 \end{vmatrix} = 0.921(-0.384) \begin{vmatrix} 0.309 & 0.265 \\ 0.492 & 0.757 \end{vmatrix} =$$

$$= 0.921(-0.384)(0.104) = -0.037$$

COFACTORES. ADJUNTOS CLASICOS

7.14. Considérese la matriz $A = \begin{pmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{pmatrix}$. Determinar el cofactor del 7 en A, o sea, A_{23}.

Tenemos

$$A_{23} = (-1)^{2+3} \begin{vmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{vmatrix} = - \begin{vmatrix} 2 & 1 & 4 \\ 4 & 0 & -3 \\ 3 & -2 & 2 \end{vmatrix} =$$

$$= -(0 - 9 - 32 - 0 - 12 - 8) = -(-61) = 61$$

El exponente $2 + 3$ proviene de que 7 ocupa la segunda fila y la tercera columna.

7.15. Considérese la matriz $B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 5 & 8 & 9 \end{pmatrix}$. Hallar: *a*) $|B|$, *b*) adj B, *c*) B^{-1} usando adj B.

a) $|B| = 27 + 20 + 16 - 15 - 32 - 18 = -2$.

b) Tomamos la traspuesta de la matriz de cofactores.

$$\text{adj } B = \begin{pmatrix} \begin{vmatrix} 3 & 4 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 5 & 9 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 5 & 8 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ 8 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 5 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 5 & 8 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} \end{pmatrix}^T = \begin{pmatrix} -5 & 2 & 1 \\ -1 & 4 & -3 \\ 1 & -2 & 1 \end{pmatrix}^T = \begin{pmatrix} -5 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -3 & 1 \end{pmatrix}$$

c) Como $|B| \neq 0$,

$$B^{-1} = \frac{1}{|B|}(\text{adj } B) = \frac{1}{-2}\begin{pmatrix} -5 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & -2 & 1 \\ -\frac{1}{2} & \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

7.16. Considérese una matriz 2-cuadrada arbitraria $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *a*) Hallar adj A. *b*) Mostrar que adj (adj A) $= A$.

a) $$\text{adj } A = \begin{pmatrix} +|d| & -|c| \\ -|b| & +|a| \end{pmatrix}^T = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

b) $$\text{adj (adj } A) = \text{adj}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} +|a| & -|-c| \\ -|-b| & +|d| \end{pmatrix}^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A.$$

DETERMINANTES Y ECUACIONES LINEALES. REGLA DE CRAMER

7.17. Resolver el sistema que sigue empleando determinantes.

$$\begin{cases} ax - 2by = c \\ 3ax - 5by = 2c \end{cases} \qquad \text{donde } ab \neq 0$$

Comenzamos por calcular $D = \begin{vmatrix} a & -2b \\ 3a & -5b \end{vmatrix} = -5ab + 6ab = ab$. Dado que $D = ab \neq 0$, el sistema tiene solución única. A continuación calculamos

$$N_x = \begin{vmatrix} c & -2b \\ 2c & -5b \end{vmatrix} = -5bc + 4bc = -bc \qquad \text{y} \qquad N_y = \begin{vmatrix} a & c \\ 3a & 2c \end{vmatrix} = 2ac - 3ac = -ac$$

Entonces $x = N_x/D = -bc/ab = -c/a$ e $y = N_y/D = -ac/ab = -c/b$.

7.18. Resolver, utilizando determinantes: $\begin{cases} 3y + 2x = z + 1 \\ 3x + 2z = 8 - 5y. \\ 3z - 1 \ \ = x - 2y \end{cases}$

Primero colocamos las ecuaciones en la forma convencional:

$$\begin{aligned} 2x + 3y - \ \ z &= 1 \\ 3x + 5y + 2z &= 8 \\ x - 2y - 3z &= -1 \end{aligned}$$

Calculamos el determinante D de la matriz de los coeficientes:

$$D = \begin{vmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{vmatrix} = -30 + 6 + 6 + 5 + 8 + 27 = 22$$

Siendo $D \neq 0$ el sistema tiene solución única. Para calcular N_x, N_y y N_z sustituimos los coeficientes de cada incógnita en la matriz de los coeficientes por los términos constantes:

$$N_x = \begin{vmatrix} 1 & 3 & -1 \\ 8 & 5 & 2 \\ -1 & -2 & -3 \end{vmatrix} = -15 - 6 + 16 - 5 + 4 + 72 = 66$$

$$N_y = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 8 & 2 \\ 1 & -1 & -3 \end{vmatrix} = -48 + 2 + 3 + 8 + 4 + 9 = -22$$

$$N_z = \begin{vmatrix} 2 & 3 & 1 \\ 3 & 5 & 8 \\ 1 & -2 & -1 \end{vmatrix} = -10 + 24 - 6 - 5 + 32 + 9 = 44$$

Por tanto,

$$x = \frac{N_x}{D} = \frac{66}{22} = 3 \qquad y = \frac{N_y}{D} = \frac{-22}{22} = -1 \qquad z = \frac{N_z}{D} = \frac{44}{22} = 2$$

7.19. Resolver el siguiente sistema mediante la regla de Cramer:

$$\begin{aligned} 2x_1 + \ \ x_2 + 5x_3 + \ \ x_4 &= \ \ 5 \\ x_1 + \ \ x_2 - 3x_3 - 4x_4 &= -1 \\ 3x_1 + 6x_2 - 2x_3 + \ \ x_4 &= \ \ 8 \\ 2x_1 + 2x_2 + 2x_3 - 3x_4 &= \ \ 2 \end{aligned}$$

Calculamos

$$D = \begin{vmatrix} 2 & 1 & 5 & 1 \\ 1 & 1 & -3 & -4 \\ 3 & 6 & -2 & 1 \\ 2 & 2 & 2 & -3 \end{vmatrix} = -120 \qquad N_1 = \begin{vmatrix} 5 & 1 & 5 & 1 \\ -1 & 1 & -3 & -4 \\ 8 & 6 & -2 & 1 \\ 2 & 2 & 2 & -3 \end{vmatrix} = -240$$

$$N_2 = \begin{vmatrix} 2 & 5 & 5 & 1 \\ 1 & -1 & -3 & -4 \\ 3 & 8 & -2 & 1 \\ 2 & 2 & 2 & -3 \end{vmatrix} = -24 \qquad N_3 = \begin{vmatrix} 2 & 1 & 5 & 1 \\ 1 & 1 & -1 & -4 \\ 3 & 6 & 8 & 1 \\ 2 & 2 & 2 & -3 \end{vmatrix} = 0$$

$$N_4 = \begin{vmatrix} 2 & 1 & 5 & 5 \\ 1 & 1 & -3 & -1 \\ 3 & 6 & -2 & 8 \\ 2 & 2 & 2 & 2 \end{vmatrix} = -96$$

Entonces $x_1 = N_1/D = 2$, $x_2 = N_2/D = \frac{1}{5}$, $x_3 = N_3/D = 0$, $x_4 = N_4/D = \frac{4}{5}$.

7.20. Utilizar determinantes para encontrar los valores de k para los que el sistema escrito a continuación tiene solución única:

$$\begin{aligned} kx + \ \ y + \ \ z &= 1 \\ x + ky + \ \ z &= 1 \\ x + \ \ y + kz &= 1 \end{aligned}$$

El sistema tiene solución única cuando $D \neq 0$, donde D es el determinante de la matriz de los coeficientes. Calculamos

$$D = \begin{vmatrix} k & 1 & 1 \\ 1 & k & 1 \\ 1 & 1 & k \end{vmatrix} = k^3 + 1 + 1 - k - k - k = k^3 - 3k + 2 = (k-1)^2(k+2)$$

Así el sistema tiene solución única cuando $(k-1)^2(k+2) \neq 0$, es decir, cuando $k \neq 1$ y $k \neq -2$. (La eliminación gaussiana indica que el sistema no tiene solución para $k = -2$, mientras que para $k = 1$ tiene un número infinito de ellas.)

DEMOSTRACION DE LOS TEOREMAS

7.21. Demostrar el Teorema 7.1.

Si $A = (a_{ij})$, $A^T = (b_{ij})$, con $b_{ij} = a_{ji}$. Por consiguiente,

$$|A^T| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) b_{1\sigma(2)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

Sea $\tau = \sigma^{-1}$. Según el Problema 7.43, $\operatorname{sgn} \tau = \operatorname{sgn} \sigma$ y $a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} = a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$. De aquí

$$|A^T| = \sum_{\sigma \in S_n} (\operatorname{sgn} \tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$$

No obstante, al recorrer σ todos los elementos de S_n, también lo hace $\tau = \sigma^{-1}$. De este modo, $|A^T| = |A|$.

7.22. Demostrar el Teorema 7.3 *a*).

Lo demostramos en el caso en que se intercambian dos columnas. Sea τ la trasposición que intercambia los números correspondientes a las dos columnas de A en cuestión. Si $A = (a_{ij})$ y $B = (b_{ij})$, necesariamente $b_{ij} = a_{i\tau(j)}$. Para cualquier permutación σ tendremos, pues,

$$b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = a_{1\tau\sigma(1)} a_{2\tau\sigma(2)} \cdots a_{n\tau\sigma(n)}$$

Así pues,

$$|B| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1\tau\sigma(1)} a_{2\tau\sigma(2)} \cdots a_{n\tau\sigma(n)}$$

Dado que la trasposición τ es una permutación impar, $\operatorname{sgn} \tau\sigma = \operatorname{sgn} \tau \cdot \operatorname{sgn} \sigma = -\operatorname{sgn} \sigma$, de modo que $\operatorname{sgn} \sigma = -\operatorname{sgn} \tau\sigma$, y

$$|B| = -\sum_{\sigma \in S_n} (\operatorname{sgn} \tau\sigma) a_{1\tau\sigma(1)} a_{2\tau\sigma(2)} \cdots a_{n\tau\sigma(n)}$$

Pero al ir recorriendo σ todos los elementos de S_n, también lo hace $\tau\sigma$, por lo que $|B| = -|A|$.

7.23. Demostrar el Teorema 7.2.

a) Cada término en $|A|$ contiene un factor de cada fila y por tanto uno de la fila de ceros. Siendo así, cada término de $|A|$ es cero y por ende $|A| = 0$.

b) Supongamos que $1 + 1 \neq 0$ en K. Si intercambiamos dos filas idénticas de A, seguiremos teniendo la matriz A, luego, de acuerdo con el Problema 7.22, $|A| = -|A|$ y por tanto $|A| = 0$.

Supongamos ahora que $1 + 1 = 0$ en K. En tal caso, $\operatorname{sgn} \sigma = 1$ para toda $\sigma \in S_n$. Como A tiene dos filas idénticas, podemos agrupar los términos de A en pares de términos iguales. Siendo nulo cada par, el determinante de A es cero.

c) Supongamos que $A = (a_{ij})$ es triangular inferior, o sea, que las entradas situadas sobre la diagonal son todas cero: $a_{ij} = 0$ siempre que $i < j$. Consideremos un término t del determinante de A:

$$t = (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n} \qquad \text{donde } \sigma = i_1 i_2 \cdots i_n$$

Supongamos $i_1 \neq 1$. Entonces $1 < i_1$, de modo que $a_{1i_1} = 0$; por consiguiente, $t = 0$. Esto es, todo término para el que $i_1 \neq 1$ es nulo.

Supongamos ahora $i_1 = 1$ pero $i_2 \neq 2$. En ese caso, $2 < i_2$ y $a_{2i_2} = 0$; por tanto, $t = 0$. Todo término para el que $i_1 \neq 1$ o $i_2 \neq 2$ es, pues, nulo.

Similarmente, obtenemos que todo término para el que $i_1 \neq 1$ o $i_2 \neq 2$ o $\cdots$ o $i_n \neq n$ es nulo. En consecuencia, $|A| = a_{11} a_{22} \cdots a_{nn}$ = producto de los elementos diagonales.

7.24. Demostrar el Teorema 7.3.

a) Demostrado en el Problema 7.22.

b) Si se multiplica por k la fila j-ésima de A, todo término en $|A|$ aparecerá multiplicado por k, de donde $|B| = k|A|$. Es decir,

$$|B| = \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots (k a_{ji_j}) \cdots a_{ni_n} = k \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n} = k|A|$$

c) Supongamos que se suma a la fila j-ésima de A c veces la k-ésima. Usando el símbolo $\wedge$ para denotar la j-ésima posición en un término del determinante.

$$|B| = \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots \widehat{(ca_{ki_k} + a_{ji_j})} \cdots a_{ni_n}$$

$$= c \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots \widehat{a_{ki_k}} \cdots a_{ni_n} + \sum_{\sigma} (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots \widehat{a_{ji_j}} \cdots s_{ni_n}$$

La primera suma es el determinante de una matriz cuyas filas k-ésima y j-ésima son idénticas; por tanto, según el Teorema 7.2 *b*), la suma es cero. La segunda suma es el determinante de A. Así pues, $|B| = c \cdot 0 + |A| = |A|$.

7.25. Demostrar el Lema 7.6.

Consideremos las operaciones elementales entre las filas siguientes: i) multiplicar una fila por una constante $k \neq 0$; ii) intercambiar dos filas; iii) sumar a una fila un múltiplo de otra. Sean E_1, E_2 y E_3 las matrices elementales asociadas, esto es, las obtenidas efectuando las operaciones precedentes, respectivamente, sobre la matriz identidad I. Por el Problema 7.24,

$$|E_1| = k|I| = k \qquad |E_2| = -|I| = -1 \qquad |E_3| = |I| = 1$$

Recordemos (Teorema 4.12) que E_iA es idéntica a la matriz obtenida efectuando sobre A la operación correspondiente. Así, según el Teorema 7.3,

$$|E_1A| = k|A| = |E_1||A| \qquad |E_2A| = -|A| = |E_2||A| \qquad |E_3A| = |A| = 1|A| = |E_3||A|$$

y queda demostrado el lema.

7.26. Supóngase que B es equivalente por filas a una matriz cuadrada A. Probar que $|B| = 0$ si y sólo si $|A| = 0$.

De acuerdo con el Teorema 7.3, el efecto de una operación elemental entre filas es cambiar el signo del determinante o multiplicar éste por un escalar no nulo. Por consiguiente, $|B| = 0$ si y sólo si $|A| = 0$.

7.27. Demostrar el Teorema 7.4.

La demostración se basa en el algoritmo de eliminación gaussiana. Si A es invertible, es equivalente por filas a I. Pero $|I| \neq 0$, luego, según el Problema 7.26, $|A| \neq 0$. Si A no es invertible, es equivalente por filas a una matriz con una fila nula, por lo que $\det(A) = 0$. Entonces i) y iii) son equivalentes.

Si $AX = 0$ sólo tiene la solución $X = 0$, A es necesariamente equivalente por filas a I y por ende invertible. Recíprocamente, si A es invertible con inversa A^{-1},

$$X = IX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}0 = 0$$

es la única solución de $AX = 0$. De este modo, i) y ii) son equivalentes.

7.28. Demostrar el Teorema 7.5.

Si A es singular, AB también lo es y $|AB| = 0 = |A||B|$. Por otra parte, si A es no singular, $A = E_n \cdots E_2E_1$, producto de matrices elementales. Haciendo uso del Lema 7.6 y de la inducción obtenemos

$$|AB| = |E_n \cdots E_2E_1B| = |E_n| \cdots |E_2||E_1||B| = |A||B|$$

7.29. Supóngase que P es invertible. Probar que $|P^{-1}| = |P|^{-1}$.

$P^{-1}P = I$. De aquí $1 = |I| = |P^{-1}P| = |P^{-1}||P|$ y así $|P^{-1}| = |P|^{-1}$.

7.30. Demostrar el Teorema 7.7.

Dado que A y B son similares, existe una matriz invertible P tal que $B = P^{-1}AP$. Por el problema precedente, $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A||P^{-1}||P| = |A|$.

Hacemos notar que aunque las matrices P^{-1} y A pueden no conmutar, sus determinantes $|P^{-1}|$ y $|A|$ sí lo hacen, por ser escalares en el cuerpo K.

7.31. Si $A = (a_{ij})$, demostrar que $|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$, donde A_{ij} es el cofactor de a_{ij}.

Cada término en $|A|$ contiene una y sólo una entrada de la i-ésima fila $(a_{i1}, a_{i2}, \ldots, a_{in})$ de A. Por tanto, podemos escribir $|A|$ de la forma

$$|A| = a_{i1}A_{i1}^* + a_{i2}A_{i2}^* + \cdots + a_{in}A_{in}^*$$

(Nótese que A_{ij}^* es una suma de términos sin entradas de la i-ésima fila de A.) El teorema quedará demostrado si probamos que

$$A_{ij}^* = A_{ij} = (-1)^{i+j}|M_{ij}|$$

siendo M_{ij} la matriz obtenida suprimiendo la fila y la columna que contienen la entrada a_{ij}. (Históricamente, la expresión A_{ij}^* se definía como el cofactor de a_{ij}, por lo que el teorema se reduce a mostrar que ambas definiciones de cofactor son equivalentes.)

Empecemos considerando el caso $i = n$, $j = n$. Entonces la suma de términos en $|A|$ que contienen a_{nn} es

$$a_{nn}A_{nn}^* = a_{nn}\sum_{\sigma}(\operatorname{sgn}\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n-1,\sigma(n-1)}$$

donde se suma sobre todas las permutaciones $\sigma \in S_n$ para las que $\sigma(n) = n$. No obstante, esto es equivalente (demuéstrese) a sumar sobre todas las permutaciones de $\{1, \ldots, n-1\}$. De este modo, $A_{nn}^* = |M_{nn}| = (-1)^{n+n}|M_{nn}|$.

Consideremos ahora unos i y j cualesquiera. Intercambiamos la fila i-ésima con cada una de las subsiguientes hasta llevarla a la última posición y hacemos lo mismo con la columna j-ésima y las subsiguientes. Adviértase que el determinante $|M_{ij}|$ no se ve afectado, puesto que los intercambios no alteran la posición relativa de las filas y columnas restantes. Sin embargo, el «signo» de $|A|$ y de A_{ij}^* varía $n-i$ y después $n-j$ veces. En consecuencia,

$$A_{ij}^* = (-1)^{n-i+n-j}|M_{ij}| = (-1)^{i+j}|M_{ij}|$$

7.32. Sean $A = (a_{ij})$ y B la matriz obtenida de A sustituyendo su fila i-ésima por el vector fila $(b_{i1}, \ldots, b_{in})$. Probar que

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

Probar además que para $j \neq i$,

$$a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0 \qquad \text{y} \qquad a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$$

Sea $B = (b_{ij})$. De acuerdo con el Teorema 7.8,

$$|B| = b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in}$$

Como B_{ij} no depende de la fila i-ésima de B, $B_{ij} = A_{ij}$ para $j = 1, \ldots, n$. Por consiguiente,

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

Sea ahora A' la matriz obtenida de A sustituyendo su fila i-ésima por la j-ésima. Como A' tiene dos filas idénticas, $|A'| = 0$. Así, según el resultado anterior,

$$|A'| = a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0$$

Utilizando $|A^T| = |A|$ llegamos también a $a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$.

7.33. Demostrar el Teorema 7.9.

Sean $A = (a_{ij})$ y $A \cdot (\text{adj } A) = (b_{ij})$. La i-ésima fila de A es

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \quad [1]$$

Siendo adj A la traspuesta de la matriz de cofactores, su j-ésima columna es la traspuesta de los cofactores de la j-ésima fila de A:

$$(A_{j1}, A_{j2}, \ldots, A_{jn})^T \quad [2]$$

Ahora bien, b_{ij}, la entrada ij de $A \cdot (\text{adj } A)$, se obtiene multiplicando [1] y [2]:

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

Por el Teorema 7.8 y el Problema 7.32,

$$b_{ij} = \begin{cases} |A| & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

De acuerdo con esto, $A \cdot (\text{adj } A)$ es la matriz diagonal con cada elemento diagonal igual a $|A|$. Dicho de otro modo, $A \cdot (\text{adj } A) = |A|I$. De forma similar, $(\text{adj } A) \cdot A = |A|I$.

7.34. Demostrar el Teorema 7.10.

Por resultados previos sabemos que $AX = B$ tiene solución única si y sólo si A es invertible, lo que se da, a su vez, si y sólo si $D = |A| \neq 0$.

Supongamos $D \neq 0$. Según el Teorema 7.9, $A^{-1} = (1/D)(\text{adj } A)$. Multiplicando $AX = B$ por A^{-1} conseguimos

$$X = A^{-1}AX = (1/D)(\text{adj } A)B \quad [1]$$

Nótese que la i-ésima fila de $(1/D)(\text{adj } A)$ es $(1/D)(A_{1i}, A_{2i}, \ldots, A_{ni})$. Si $B = (b_1, b_2, \ldots, b_n)^T$, tendremos, por [1],

$$x_i = (1/D)(b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni})$$

No obstante, como en el Problema 7.32, $b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni} = N_i$, que es el determinante de la matriz obtenida sustituyendo la i-ésima fila de A por el vector columna B. Así $x_i = (1/D)N_i$, como se pedía.

7.35. Demostrar el Teorema 7.12.

Sólo es necesario probar el teorema para $n = 2$, esto es, cuando M es una matriz cuadrada por bloques de la forma $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$. La demostración general se deriva fácilmente por inducción.

Supongamos que $A = (a_{ij})$ es r-cuadrada, $B = (b_{ij})$, s-cuadrada y $M = (m_{ij})$, n-cuadrada, con $n = r + s$. Por definición,

$$\det M = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)}$$

Si $i > r$ y $j \leq r$, entonces $m_{ij} = 0$, de modo que sólo es necesario considerar aquellas permutaciones σ tales que

$$\sigma\{r+1, r+2, \ldots, r+s\} = \{r+1, r+2, \ldots, r+s\} \qquad \text{y} \qquad \sigma\{1, 2, \ldots, r\} = \{1, 2, \ldots, r\}$$

Sean $\sigma_1(k) = \sigma(k)$ para $k \leq r$ y $\sigma_2(k) = \sigma(r+k) - r$ para $k \leq s$. Entonces

$$(\operatorname{sgn} \sigma) m_{1\sigma(1)} m_{2\sigma(2)} \cdots m_{n\sigma(n)} = (\operatorname{sgn} \sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} \cdots a_{r\sigma_2(r)} (\operatorname{sgn} \sigma_2) b_{1\sigma_2(1)} b_{2\sigma_2(2)} \cdots b_{s\sigma_2(s)}$$

lo que implica $\det M = (\det A)(\det B)$.

7.36. Demostrar el Teorema 7.13.

Sea D la función determinante: $D(A) = |A|$. Debemos probar que D satisface i), ii) y iii) y que es la única función que lo hace.

De acuerdo con el Teorema 7.2, D satisface ii) y iii); por consiguiente, sólo necesitamos mostrar que es multilineal. Supongamos que la fila i-ésima de $A = (a_{ij})$ es de la forma $(b_{i1} + c_{i1}, b_{i2} + c_{i2}, \ldots, b_{in} + c_{in})$. En tal caso,

$$\begin{aligned} D(A) &= D(A_1, \ldots, B_i + C_i, \ldots, A_n) = \\ &= \sum_{S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \cdots a_{i-1, \sigma(i-1)} (b_{i\sigma(i)} + c_{i\sigma(i)}) \cdots a_{n\sigma(n)} = \\ &= \sum_{S_n} (\operatorname{sgn} \sigma) a_{i\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)} = \\ &= D(A_1, \ldots, B_i, \ldots, A_n) + D(A_1, \ldots, C_i, \ldots, A_n) \end{aligned}$$

Asimismo, según el Teorema 7.3 *b*),

$$D(A_1, \ldots, kA_i, \ldots, A_n) = kD(A_1, \ldots, A_i, \ldots, A_n)$$

Así pues, D es multilineal, es decir, satisface i).

Probemos ahora la unicidad de D. Supongamos que D satisface i), ii) y iii). Si $\{e_1, \ldots, e_n\}$ es la base usual de K^n, por iii), $D(e_1, e_2, \ldots, e_n) = D(I) = 1$. Usando ii) tenemos también

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = \operatorname{sgn} \sigma \qquad \text{donde } \sigma = i_1 i_2 \cdots i_n \qquad [1]$$

Supongamos ahora $A = (a_{ij})$. Observamos que la fila k-ésima A_k de A es

$$A_k = (a_{k1}, a_{k2}, \ldots, a_{kn}) = a_{k1} e_1 + a_{k2} e_2 + \cdots + a_{kn} e_n$$

de modo que

$$D(A) = D(a_{11} e_1 + \cdots + a_{1n} e_n, a_{21} e_1 + \cdots + a_{2n} e_n, \ldots, a_{n1} e_1 + \cdots + a_{nn} e_n)$$

Utilizando la multilinealidad de D podemos escribir $D(A)$ como una suma de términos de la forma

$$\begin{aligned} D(A) &= \sum D(a_{1i_1} e_{i_1}, a_{2i_2} e_{i_2}, \ldots, a_{ni_n} e_{i_n}) = \\ &= \sum (a_{1i_1} a_{2i_2} \cdots a_{ni_n}) D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \end{aligned} \qquad [2]$$

donde la suma se extiende a todas las sucesiones $i_1 i_2 \cdots i_n$ con $i_k \in \{1, \ldots, n\}$. Si dos de los índices son iguales, digamos $i_j = i_k$ pero $j \neq k$, por ii),

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = 0$$

En consecuencia, sólo es necesario efectuar la suma en [2] sobre todas las permutaciones $\sigma = i_1 i_2 \cdots i_n$. Haciendo uso de [1] tenemos, finalmente,

$$D(A) = \sum_\sigma (a_{1i_1} a_{2i_2} \cdots a_{ni_n}) D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) =$$
$$= \sum_\sigma (\operatorname{sgn} \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n} \qquad \text{donde } \sigma = i_1 i_2 \cdots i_n$$

Por tanto, D es la función determinante y queda así demostrado el teorema.

PERMUTACIONES

7.37. Determinar la paridad de $\sigma = 542163$.

Método 1. Queremos obtener el número de pares (i, j) para los que $i > j$ e i precede a j en σ. Hay:

3 números (5, 4 y 2) mayores que 1 y precediéndolo,
2 números (5 y 4) mayores que 2 y precediéndolo,
3 números (5, 4 y 6) mayores que 3 y precediéndolo,
1 número (5) mayor que 4 y precediéndolo,
0 números mayores que 5 y precediéndolo,
0 números mayores que 6 y precediéndolo.

Como $3 + 2 + 3 + 1 + 0 + 0 = 9$ es impar, σ es una permutación impar y así $\operatorname{sgn} \sigma = -1$.

Método 2. Llevamos 1 a la primera posición como sigue:

5 4 2 (1) 6 3 a 1 5 4 2 6 3

Llevamos 2 a la segunda posición:

1 5 4 (2) 6 3 a 1 2 5 4 6 3

Llevamos 3 a la tercera posición:

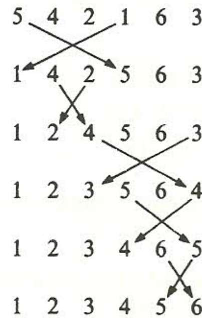
1 2 5 4 6 (3) a 1 2 3 5 4 6

Llevamos 4 a la cuarta posición:

1 2 3 5 (4) 6 a 1 2 3 4 5 6

Nótese que 5 y 6 están en las posiciones «correctas». Contamos los números sobre los que se ha «saltado»: $3 + 2 + 3 + 1 = 9$. Dado que 9 es impar, σ es una permutación impar. (*Nota:* Este método coincide esencialmente con el anterior.)

Método 3. Un intercambio de dos números en una permutación es equivalente a mutiplicar ésta por una trasposición. Transformemos, pues, σ en la permutación identidad por medio de trasposiciones, tales como,



Habiéndose empleado un número impar, 5, de trasposiciones (y puesto que impar × impar = impar), σ es una permutación impar.

7.38. Sean $\sigma = 24513$ y $\tau = 41352$ permutaciones en S_5. Hallar:

a) La permutación compuesta $\tau \circ \sigma$, *b*) $\sigma \circ \tau$, *c*) σ^{-1}.

Recordemos que $\sigma = 24513$ y $\tau = 41352$ son abreviaturas de

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix} \quad \text{y} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$

lo que significa

$$\sigma(1) = 2 \qquad \sigma(2) = 4 \qquad \sigma(3) = 5 \qquad \sigma(4) = 1 \qquad \text{y} \qquad \sigma(5) = 3$$

y

$$\tau(1) = 4 \qquad \tau(2) = 1 \qquad \tau(3) = 3 \qquad \tau(4) = 5 \qquad \text{y} \qquad \tau(5) = 2$$

a) El efecto de σ seguida de τ sobre 1, 2, ..., 5 es:

$$\begin{array}{cccccc} & 1 & 2 & 3 & 4 & 5 \\ \sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 2 & 4 & 5 & 1 & 3 \\ \tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 1 & 5 & 2 & 4 & 3 \end{array}$$

Siendo así, $\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 2 & 4 & 3 \end{pmatrix}$, o $\tau \circ \sigma = 15243$.

b) El efecto de τ seguida de σ sobre 1, 2, ..., 5 es:

$$\begin{array}{cccccc} & 1 & 2 & 3 & 4 & 5 \\ \tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 4 & 1 & 3 & 5 & 2 \\ \sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 5 & 3 & 4 \end{array}$$

Por tanto, $\sigma \circ \tau = 12534$.

c) Por definición, $\sigma^{-1}(j) = k$ si y sólo si $\sigma(k) = j$, luego

$$\sigma^{-1} = \begin{pmatrix} 2 & 4 & 5 & 1 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix} \quad \text{o} \quad \sigma^{-1} = 41523$$

7.39. Considérese cualquier permutación $\sigma = j_1 j_2 \cdots j_n$. Probar que para cada inversión (i, k) en σ existe una pareja (i^*, k^*) tal que

$$i^* < k^* \quad \text{y} \quad \sigma(i^*) > \sigma(k^*) \qquad [1]$$

y viceversa. De este modo, σ es par o impar según haya un número par o impar de parejas que satisfagan [1].

Elegimos i^* y k^* de forma que $\sigma(i^*) = i$ y $\sigma(k^*) = k$. En tal caso, $i > k$ si y sólo si $\sigma(i^*) > \sigma(k^*)$, e i precede a k en σ si y sólo si $i^* < k^*$.

7.40. Considérese el polinomio $g = g(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j)$. Escribir explícitamente $g = g(x_1, x_2, x_3, x_4)$.

El símbolo $\prod$ se emplea para un producto de términos de la misma manera que el símbolo $\sum$ se usa para una suma de términos. Esto es, $\prod_{i<j} (x_i - x_j)$ significa el producto de todos los términos $(x_i - x_j)$ para los que $i < j$. Por consiguiente,

$$g = g(x_1, \ldots, x_4) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)$$

7.41. Sea σ una permutación arbitraria. Para el polinomio precedente g del Problema 7.40, defínase $\sigma(g) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)})$. Probar que

$$\sigma(g) = \begin{cases} g & \text{si } \sigma \text{ es par} \\ -g & \text{si } \sigma \text{ es impar} \end{cases}$$

De acuerdo con ello, $\sigma(g) = (\text{sgn } \sigma)g$.

Como σ es biyectiva,

$$\sigma(g) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)}) = \prod_{i<j \text{ o } i>j} (x_i - x_j)$$

Siendo así, $\sigma(g) = g$ o $\sigma(g) = -g$ según haya un número par o impar de términos de la forma $(x_i - x_j)$ con $i > j$. Nótese que por cada pareja (i, j) para la que

$$i < j \quad \text{y} \quad \sigma(i) > \sigma(j) \qquad [1]$$

hay un término $(x_{\sigma(i)} - x_{\sigma(j)})$ en $\sigma(g)$ para el cual $\sigma(i) > \sigma(j)$. Dado que σ es par si y sólo si existe un número par de parejas que satisfacen [1], tenemos $\sigma(g) = g$ si y sólo si σ es par; de aquí $\sigma(g) = -g$ si y sólo si σ es impar.

7.42. Sean $\sigma, \tau \in S_n$. Probar que $\text{sgn } (\tau \circ \sigma) = (\text{sgn } \tau)(\text{sgn } \sigma)$. De este modo, el producto de dos permutaciones pares o dos impares es par, mientras que el producto de una permutación par y una impar es impar.

Haciendo uso del Problema 7.41 tenemos

$$\operatorname{sgn}(\tau \circ \sigma)g = (\tau \circ \sigma)(g) = \tau(\sigma(g)) = \tau((\operatorname{sgn} \sigma)g) = (\operatorname{sgn} \tau)(\operatorname{sgn} \sigma)g$$

De acuerdo con esto, $\operatorname{sgn}(\tau \circ \sigma) = (\operatorname{sgn} \tau)(\operatorname{sgn} \sigma)$.

7.43. Considérese la permutación $\sigma = j_1 j_2 \cdots j_n$. Probar que $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$ y que, para escalares a_{ij},

$$a_{j_1 1} + a_{j_2 2} \cdots a_{j_n n} = a_{1k_1} + a_{2k_2} \cdots a_{nk_n}$$

donde $\sigma^{-1} = k_1 k_2 \cdots k_n$.

Sabemos que $\sigma^{-1} \circ \sigma = \varepsilon$, la permutación identidad. Siendo ε par, σ^{-1} y σ son ambas pares o ambas impares. Por consiguiente, $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$.

Como $\sigma = j_1 j_2 \cdots j_n$ es una permutación, $a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} = a_{1k_1} a_{2k_2} \cdots a_{nk_n}$. Entonces $k_1, k_2, \ldots, k_n$ tienen la propiedad de que

$$\sigma(k_1) = 1,\ \sigma(k_2) = 2, \ldots, \sigma(k_n) = n$$

Sea $\tau = k_1 k_2 \cdots k_n$. Entonces para $i = 1, \ldots, n$,

$$(\sigma \circ \tau)(i) = \sigma(\tau(i)) = \sigma(k_i) = i$$

Así $\sigma \circ \tau = \varepsilon$, la permutación identidad, luego $\tau = \sigma^{-1}$.

PROBLEMAS VARIOS

7.44. Sin desarrollar el determinante, mostrar que $\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} = 0.$

Sumamos la segunda columna a la tercera y sacamos el factor común de esta última, lo que nos conduce a

$$\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} = \begin{vmatrix} 1 & a & a+b+c \\ 1 & b & a+b+c \\ 1 & c & a+b+c \end{vmatrix} = (a+b+c)\begin{vmatrix} 1 & a & 1 \\ 1 & b & 1 \\ 1 & c & 1 \end{vmatrix} = (a+b+c)(0) = 0$$

(Utilizamos el hecho de que el determinante de una matriz con dos columnas idénticas es nulo.)

7.45. Probar que el producto de diferencias $g(x_1, \ldots, x_n)$ del Problema 7.40 puede representarse mediante el *determinante de Vandermonde* de $x_1, x_2, \ldots, x_{n-1}, x$ definido según:

$$V_{n-1}(x) \equiv \begin{vmatrix} 1 & 1 & \ldots & 1 & 1 \\ x_1 & x_2 & \ldots & x_{n-1} & x \\ x_1^2 & x_2^2 & \ldots & x_{n-1}^2 & x^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_{n-1}^{n-1} & x^{n-1} \end{vmatrix}$$

Este es un polinomio en x de grado $n-1$, cuyas raíces son $x_1, x_2, \ldots, x_{n-1}$; más aún, el coeficiente dominante (el cofactor de x^{n-1}) es igual a $V_{n-2}(x_{n-1})$. Tendremos, pues, según un resultado del álgebra,

$$V_{n-1}(x) = (x - x_1)(x - x_2) \cdots (x - x_{n-1}) V_{n-2}(x_{n-1})$$

de modo que, por recurrencia,

$$\begin{aligned} V_{n-1}(x) &= [(x - x_1) \cdots (x - x_{n-1})][(x_{n-1} - x_1) \cdots (x_{n-1} - x_{n-2})]V_{n-3}(x_{n-2}) \\ &= \cdots \\ &= [(x - x_1) \cdots (x - x_{n-1})][(x_{n-1} - x_1) \cdots (x_{n-1} - x_{n-2})] \cdots [(x_2 - x_1)] \end{aligned}$$

De lo anterior se sigue que

$$V_{n-1}(x_n) = \prod_{n \geq i > j \geq 1} (x_i - x_j) = (-1)^{n(n-1)/2} \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

Así pues, $g(x_1, \ldots, x_n) = (-1)^{n(n-1)/2} V_{n-1}(x_n)$.

7.46. Hallar el volumen $V(S)$ del paralelepípedo S en $\mathbf{R}^3$ determinado por los vectores $u_1 = (1, 2, 4)$, $u_2 = (2, 1, -3)$ y $u_3 = (5, 7, 9)$.

Evaluamos el determinante $\begin{vmatrix} 1 & 2 & 4 \\ 2 & 1 & -3 \\ 5 & 7 & 9 \end{vmatrix} = 9 - 30 + 56 - 20 + 21 - 36 = 0$. De este modo, $V(S) = 0$ o, en otras palabras, u_1, u_2 y u_3 yacen en un plano.

7.47. Hallar el volumen $V(S)$ del paralelepípedo S en $\mathbf{R}^4$ determinado por los vectores $u_1 = (2, -1, 4, -3)$, $u_2 = (-1, 1, 0, 2)$, $u_3 = (3, 2, 3, -1)$ y $u_4 = (1, -2, 2, 3)$.

Evaluamos el siguiente determinante, usando u_{22} como pivote y efectuando $C_2 + C_1 \to C_1$ y $-2C_2 + C_4 \to C_4$:

$$\begin{vmatrix} 2 & -1 & 4 & -3 \\ -1 & 1 & 0 & 2 \\ 3 & 2 & 3 & -1 \\ 1 & -2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 4 & -1 \\ 0 & 1 & 0 & 0 \\ 5 & 2 & 3 & -5 \\ -1 & -2 & 2 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 4 & -1 \\ 5 & 3 & -5 \\ -1 & 2 & 7 \end{vmatrix} =$$

$$= 21 + 20 - 10 - 3 + 10 - 140 = -102$$

De aquí $V(S) = 102$.

7.48. Calcular det (M), donde $M = \begin{pmatrix} 3 & 4 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 0 & 9 & 2 & 0 & 0 \\ 0 & 5 & 0 & 6 & 7 \\ 0 & 0 & 4 & 3 & 4 \end{pmatrix}$.

Partimos M en una matriz triangular (inferior) por bloques como sigue:

$$M = \left(\begin{array}{cc|c|cc} 3 & 4 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ \hline 0 & 9 & 2 & 0 & 0 \\ \hline 0 & 5 & 0 & 6 & 7 \\ 0 & 0 & 4 & 3 & 4 \end{array}\right)$$

Evaluamos el determinante de cada bloque diagonal:

$$\begin{vmatrix} 3 & 4 \\ 2 & 5 \end{vmatrix} = 15 - 8 = 7 \qquad |2| = 2 \qquad \begin{vmatrix} 6 & 7 \\ 3 & 4 \end{vmatrix} = 24 - 21 = 3$$

Por consiguiente, $|M| = 7 \cdot 2 \cdot 3 = 42$.

7.49. Encontrar el menor, menor con signo y menor complementario de $A^{2,\,3}_{1,\,3}$, siendo

$$A = \begin{pmatrix} 1 & 0 & 4 & -1 \\ 3 & 2 & -2 & 5 \\ 0 & 1 & -3 & 7 \\ 6 & -4 & -5 & -1 \end{pmatrix}$$

Los subíndices de fila son 1 y 3, y los de columna, 2 y 3; por tanto, el menor es

$$|A^{2,\,3}_{1,\,3}| = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 0 & 4 \\ 1 & -3 \end{vmatrix} = 0 - 4 = -4$$

y el menor con signo,

$$(-1)^{1+3+2+3}|A^{2,\,3}_{1,\,3}| = -(-4) = 4$$

Los subíndices de fila que faltan son 2 y 4, y los de columna, 1 y 4. El menor complementario es, pues,

$$|A^{1,\,4}_{2,\,4}| = \begin{vmatrix} a_{21} & a_{24} \\ a_{41} & a_{44} \end{vmatrix} = \begin{vmatrix} 3 & 5 \\ 6 & -1 \end{vmatrix} = -3 - 30 = -33$$

7.50. Sea $A = (a_{ij})$ una matriz 3-cuadrada. Describir la suma S_k de los menores principales de órdenes *a*) $k = 1$, *b*) $k = 2$, *c*) $k = 3$.

a) Los menores principales de orden uno son los elementos diagonales, luego $S_1 = a_{11} + a_{22} + a_{33} = \operatorname{tr} A$, la traza de A.

b) Los menores principales de orden dos son los cofactores de los elementos diagonales. Así pues, $S_2 = A_{11} + A_{22} + A_{33}$, siendo A_{ii} el cofactor de a_{ii}.

c) Hay únicamente un menor principal de orden tres, el determinante de A. De este modo, $S_3 = \det(A)$.

7.51. Hallar el número N_k y la suma S_k de todos los menores principales de orden *a*) $k = 1$, *b*) $k = 2$, *c*) $k = 3$ y *d*) $k = 4$ de la matriz

$$A = \begin{pmatrix} 1 & 3 & 0 & -1 \\ -4 & 2 & 5 & 1 \\ 1 & 0 & 3 & -2 \\ 3 & -2 & 1 & 4 \end{pmatrix}$$

Cada subconjunto (no vacío) de la diagonal (o, equivalentemente, cada subconjunto no vacío de $\{1, 2, 3, 4\}$) determina un menor principal de A, y $N_k = \binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ de ellos son de orden k.

a) $N_1 = \binom{4}{1} = 4$ y

$$S_1 = |1| + |2| + |3| + |4| = 1 + 2 + 3 + 4 = 10$$

b) $N_2 = \binom{4}{2} = 6$ y

$$S_2 = \begin{vmatrix} 1 & 3 \\ -4 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 3 & 4 \end{vmatrix} + \begin{vmatrix} 2 & 5 \\ 0 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ -2 & 4 \end{vmatrix} + \begin{vmatrix} 3 & -2 \\ 1 & 4 \end{vmatrix} =$$
$$= 14 + 3 + 7 + 6 + 10 + 14 = 54$$

c) $N_3 = \binom{4}{3} = 4$ y

$$S_3 = \begin{vmatrix} 1 & 3 & 0 \\ -4 & 2 & 5 \\ 1 & 0 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 3 & -1 \\ -4 & 2 & 1 \\ 3 & -2 & 4 \end{vmatrix} + \begin{vmatrix} 1 & 0 & -1 \\ 1 & 3 & -2 \\ 3 & 1 & 4 \end{vmatrix} + \begin{vmatrix} 2 & 5 & 1 \\ 0 & 3 & -2 \\ -2 & 1 & 4 \end{vmatrix} =$$
$$= 57 + 65 + 22 + 54 = 198$$

d) $N_4 = 1$ y $S_4 = \det(A) = 378$.

7.52. Sea V el espacio vectorial de las matrices 2 por 2 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sobre **R**. Determinar si $D: V \to \mathbf{R}$ es o no 2-lineal (con respecto a las filas), donde: *a*) $D(M) = a + d$, *b*) $D(M) = ad$.

a) No. Por ejemplo, supongamos $A = (1, 1)$ y $B = (3, 3)$. Entonces

$$D(A, B) = D\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} = 4 \qquad \text{y} \qquad D(2A, B) = D\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} = 5 \neq 2D(A, B)$$

b) Sí. Sean $A = (a_1, a_2)$, $B = (b_1, b_2)$ y $C = (c_1, c_2)$; en ese caso,

$$D(A, C) = D\begin{pmatrix} a_1 & a_2 \\ c_1 & c_2 \end{pmatrix} = a_1c_2 \qquad \text{y} \qquad D(B, C) = D\begin{pmatrix} b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} = b_1c_2$$

Por tanto, para todo par de escalares $s, t \in \mathbf{R}$,

$$D(sA + tB, C) = D\begin{pmatrix} sa_1 + tb_1 & sa_2 + tb_2 \\ c_1 & c_2 \end{pmatrix} = (sa_1 + tb_1)c_2 =$$
$$= s(a_1c_2) + t(b_1c_2) = sD(A, C) + tD(B, C)$$

O sea, D es lineal con respecto a la primera fila.
Además,

$$D(C, A) = D\begin{pmatrix} c_1 & c_2 \\ a_1 & a_2 \end{pmatrix} = c_1a_2 \qquad \text{y} \qquad D(C, B) = D\begin{pmatrix} c_1 & c_2 \\ b_1 & b_2 \end{pmatrix} = c_1b_2$$

luego para todo par de escalares s, $t \in \mathbf{R}$,

$$D(C, aA + tB) = D\begin{pmatrix} c_1 & c_2 \\ sa_1 + tb_1 & sa_2 + tb_2 \end{pmatrix} = c_1(sa_2 + tb_2) =$$
$$= s(c_1 a_2) + t(c_1 b_2) = sD(C, A) + tD(C, B)$$

Es decir, D es lineal con respecto a la segunda fila.

Las dos condiciones de linealidad implican que D es 2-lineal.

7.53. Sea D una función 2-lineal, alternada. Mostrar que $D(A, B) = -D(B, A)$. Con mayor generalidad, probar que si D es multilineal y alternada,

$$D(\ldots, A, \ldots, B, \ldots) = -D(\ldots, B, \ldots, A, \ldots)$$

esto es, el signo varía cada vez que se intercambian dos variables.

Siendo D alternada, $D(A + B, A + B) = 0$. Además, como D es multilineal,

$$0 = D(A + B, A + B) = D(A, A + B) + D(B, A + B) =$$
$$= D(A, A) + D(A, B) + D(B, A) + D(B, B)$$

Pero $D(A, A) = 0$ y $D(B, B) = 0$. De aquí

$$0 = D(A, B) + D(B, A) \qquad \text{o} \qquad D(A, B) = -D(B, A)$$

De forma similar,

$$0 = D(\ldots, A + B, \ldots, A + B, \ldots) =$$
$$= D(\ldots, A, \ldots, A, \ldots) + D(\ldots, A, \ldots, B, \ldots) + D(\ldots, B, \ldots, A, \ldots) + D(\ldots, B, \ldots, B, \ldots) =$$
$$= D(\ldots, A, \ldots, B, \ldots) + D(\ldots, B, \ldots, A, \ldots)$$

y así $D(\ldots, A, \ldots, B, \ldots) = -D(\ldots, B, \ldots, A, \ldots)$.

PROBLEMAS SUPLEMENTARIOS

CALCULO DE DETERMINANTES

7.54. Calcular el determinante de cada matriz:

$$a) \begin{pmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{pmatrix}, \quad b) \begin{pmatrix} 3 & -2 & -4 \\ 2 & 5 & -1 \\ 0 & 6 & 1 \end{pmatrix}, \quad c) \begin{pmatrix} -2 & -1 & 4 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{pmatrix}, \quad d) \begin{pmatrix} 7 & 6 & 5 \\ 1 & 2 & 1 \\ 3 & -2 & 1 \end{pmatrix}$$

7.55. Evaluar el determinante de cada matriz:

$$a) \begin{pmatrix} t-2 & 4 & 3 \\ 1 & t+1 & -2 \\ 0 & 0 & t-4 \end{pmatrix}, \quad b) \begin{pmatrix} t-1 & 3 & -3 \\ -3 & t+5 & -3 \\ -6 & 6 & y-4 \end{pmatrix}, \quad c) \begin{pmatrix} t+3 & -1 & 1 \\ 7 & t-5 & 1 \\ 6 & -6 & t+2 \end{pmatrix}$$

7.56. Para cada una de las matrices del Problema 7.55, determinar aquellos valores de t para los que el determinante es nulo.

7.57. Evaluar el determinante de cada matriz: *a*) $\begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 3 & -1 & 1 & -2 \\ 4 & -3 & 0 & 2 \end{pmatrix}$, *b*) $\begin{pmatrix} 2 & 1 & 3 & 2 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{pmatrix}$.

7.58. Evaluar cada uno de los determinantes:

a) $\begin{vmatrix} 1 & 2 & -1 & 3 & 1 \\ 2 & -1 & 1 & -2 & 3 \\ 3 & 1 & 0 & 2 & -1 \\ 5 & 1 & 2 & -3 & 4 \\ -2 & 3 & -1 & 1 & -2 \end{vmatrix}$, *b*) $\begin{vmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 6 & 2 \\ 0 & 0 & 2 & 3 & 1 \end{vmatrix}$, *c*) $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 6 & 5 & 1 \\ 0 & 0 & 0 & 7 & 4 \\ 0 & 0 & 0 & 2 & 3 \end{vmatrix}$

COFACTORES. ADJUNTOS CLASICOS. INVERSAS

7.59. Sea $A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Hallar: *a*) adj A, *b*) A^{-1}.

7.60. Hallar el adjunto clásico de cada una de las matrices del Problema 7.57.

7.61. Determinar la matriz 2 por 2 general A para la que $A = \text{adj } A$.

7.62. Supóngase que A es diagonal y B triangular; por ejemplo,

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} b_1 & c_{12} & \dots & c_{1n} \\ 0 & b_2 & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_n \end{pmatrix}$$

a) Probar que adj A es diagonal y adj B triangular.

b) Mostrar que B es invertible si y sólo si todos los $b_i \neq 0$, de modo que A es invertible si y sólo si todos los $a_i \neq 0$.

c) Probar que las inversas de A y B (si existen) son de la forma

$$A^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \dots & 0 \\ 0 & a_2^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n^{-1} \end{pmatrix}, \qquad B^{-1} = \begin{pmatrix} b_1^{-1} & d_{12} & \dots & d_{1n} \\ 0 & b_2^{-1} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_n^{-1} \end{pmatrix}$$

O sea, los elementos diagonales de A^{-1} y B^{-1} son los inversos de los elementos diagonales correspondientes de A y B.

DETERMINANTES Y ECUACIONES LINEALES

7.63. Resolver por determinantes: *a*) $\begin{cases} 3x + 5y = 8 \\ 4x - 2y = 1 \end{cases}$, *b*) $\begin{cases} 2x - 3y = -1 \\ 4x + 7y = -1 \end{cases}$.

7.64. Resolver por determinantes: *a*) $\begin{cases} 2x - 5y + 2z = 7 \\ x + 2y - 4z = 3 \\ 3x - 4y - 6z = 5 \end{cases}$, *b*) $\begin{cases} 2z + 3 = y + 3x \\ x - 3z = 2y + 1 \\ 3y + z = 2 - 2x \end{cases}$.

7.65. Demostrar el Teorema 7.11.

PERMUTACIONES

7.66. Determinar la paridad de las permutaciones de S_5: *a*) $\sigma = 32154$, *b*) $\tau = 13524$, *c*) $\pi = 42531$.

7.67. Para las permutaciones σ, τ y π del Problema 7.66, hallar: *a*) $\tau \circ \sigma$, *b*) $\pi \circ \sigma$, *c*) σ^{-1}, *d*) τ^{-1}.

7.68. Sea $\tau \in S_n$. Probar que $\tau \circ \sigma$ recorre S_n a medida que lo hace σ; es decir, $S_n = \{\tau \circ \sigma : \sigma \in S_n\}$.

7.69. Supóngase que $\sigma \in S_n$ tiene la propiedad de que $\sigma(n) = n$. Defínase $\sigma^* \in S_{n-1}$ por $\sigma^*(x) = \sigma(x)$. *a*) Probar que sgn $\sigma^* =$ sgn σ. *b*) Probar que al recorrer σ S_n, con $\sigma(n) = n$, σ^* recorre S_{n-1}; esto es, $S_{n-1} = \{\sigma^* : \sigma \in S_n, \sigma(n) = n\}$.

7.70. Considérese una permutación $\sigma = j_1 j_2 \cdots j_n$. Sean $\{e_i\}$ la base usual de K^n y A la matriz cuya fila i-ésima es e_{j_i}, por ejemplo, $A = (e_{j_1}, e_{j_2}, \ldots, e_{j_n})$. Mostrar que $|A| = \text{sgn}\ \sigma$.

PROBLEMAS VARIOS

7.71. Hallar el volumen $V(S)$ del paralelepípedo S en $\mathbf{R}^3$ determinado por los vectores:

a) $u_1 = (1, 2, -3)$, $u_2 = (3, 4, -1)$, $u_3 = (2, -1, 5)$; *b*) $u_1 = (1, 1, 3)$, $u_2 = (1, -2, -4)$, $u_3 = (4, 1, 2)$.

7.72. Hallar el volumen $V(S)$ del paralelepípedo S en $\mathbf{R}^4$ determinado por los vectores:

$u_1 = (1, -2, 5, -1)$ $\quad u_2 = (2, 1, -2, 1)$ $\quad u_3 = (3, 0, 1, -2)$ $\quad u_4 = (1, -1, 4, -1)$

7.73. Encontrar el menor M_1, el menor con signo M_2 y el menor complementario M_3 de $A^{3;\,4}_{1;\,4}$, donde:

a) $A = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 1 & 0 & -2 & 3 \\ 3 & -1 & 2 & 5 \\ 4 & -3 & 0 & -1 \end{pmatrix}$, *b*) $A = \begin{pmatrix} 1 & 3 & -1 & 5 \\ 2 & -3 & 1 & 4 \\ 0 & -5 & 2 & 1 \\ 3 & 0 & 5 & -2 \end{pmatrix}$

7.74. Para $k = 1, 2, 3$, hallar la suma S_k de los menores principales de orden k de:

a) $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & -4 & 3 \\ 5 & -2 & 1 \end{pmatrix}$, *b*) $B = \begin{pmatrix} 1 & 5 & -4 \\ 2 & 6 & 1 \\ 3 & -2 & 0 \end{pmatrix}$, *c*) $C = \begin{pmatrix} 1 & -4 & 3 \\ 2 & 1 & 5 \\ 4 & -7 & 11 \end{pmatrix}$

7.75. Para $k = 1, 2, 3, 4$, hallar la suma S_k de los menores principales de orden k de:

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ 1 & -2 & 0 & 5 \\ 0 & 1 & -2 & 2 \\ 4 & 0 & -1 & -3 \end{pmatrix}$$

7.76. Sea A una matriz n-cuadrada. Demostrar que $|kA| = k^n|A|$.

7.77. Sean A, B, C y D matrices n-cuadradas que conmutan. Considérese la matriz por bloques $2n$-cuadrada $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Demostrar que $|M| = |A|\,|D| - |B|\,|C|$. Probar que el resultado puede no ser cierto si las matrices no conmutan.

7.78. Supóngase que A es ortogonal, es decir, $A^T A = I$. Probar que $\det(A) = \pm 1$.

7.79. Sea V el espacio de las matrices 2×2 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sobre $\mathbf{R}$. Determinar si $D: V \to \mathbf{R}$ es o no 2-lineal (con respecto a las filas), donde: *a*) $D(M) = ac - bd$, *b*) $D(M) = ab - cd$, *c*) $D(M) = 0$, *d*) $D(M) = 1$.

7.80. Sea V el espacio de las matrices m-cuadradas vistas como m-plas de vectores fila. Supóngase que $D: V \to K$ es m-lineal y alternada. Mostrar que si $A_1, A_2, \ldots, A_m$ son linealmente dependientes, necesariamente $D(A_1, \ldots, A_m) = 0$.

7.81. Sea V el espacio de las matrices m-cuadradas (como arriba) y supóngase $D: V \to K$. Probar que la siguiente condición más débil equivale a que D sea alternada:

$$D(A_1, A_2, \ldots, A_n) = 0 \qquad \text{siempre que } A_i = A_{i+1} \text{ para algún } i$$

7.82. Sea V el espacio de las matrices n-cuadradas sobre K. Supóngase que $B \in V$ es invertible y por tanto $\det(B) \neq 0$. Defínase $D: V \to K$ por $D(A) = \det(AB)/\det(B)$, donde $A \in V$, de modo que

$$D(A_1, A_2, \ldots, A_n) = \det(A_1B, A_2B, \ldots, A_nB)/\det(B)$$

donde A_i es la fila i-ésima de A y así A_iB es la de AB. Probar que D es multilineal y alternada y que $D(I) = 1$. (Este método se utiliza en algunos textos para demostrar que $|AB| = |A|\,|B|$.)

7.83. Sea A una matriz n-cuadrada. El *rango por determinantes* de A es el orden de la mayor submatriz de A (obtenida suprimiendo filas y columnas de la misma) cuyo determinante no es nulo. Mostrar que el rango por determinantes de A es igual a su rango, o sea, al número máximo de filas (o columnas) linealmente independientes.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

7.54. *a*) 21, *b*) -11, *c*) 100, *d*) 0.

7.55. *a*) $(t + 2)(t - 3)(t - 4)$, *b*) $(t + 2)^2(t - 4)$, *c*) $(t + 2)^2(t - 4)$.

7.56. *a*) $3, 4, -2$; *b*) $4, -2$; *c*) $4, -2$.

7.57. *a*) -131, *b*) -55.

7.58. *a*) -12, *b*) -42, *c*) -468.

7.59. $\operatorname{adj} A = \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix}$

7.60. *a*) $\begin{pmatrix} -16 & -29 & -26 & -2 \\ -30 & -38 & -16 & 29 \\ -8 & 51 & -13 & -1 \\ -13 & 1 & 28 & -18 \end{pmatrix}$, *b*) $\begin{pmatrix} 21 & -14 & -17 & -19 \\ -44 & 11 & 33 & 11 \\ -29 & 1 & 13 & 21 \\ 17 & 7 & -19 & -18 \end{pmatrix}$

7.61. $A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$

7.63. *a*) $x = \frac{21}{26}, y = \frac{29}{26}$; *b*) $x = -\frac{5}{13}, y = \frac{1}{13}$.

7.64. *a*) $x = 5, y = 1, z = 1$. *b*) Como $D = 0$, el sistema no puede resolverse por determinantes.

7.66. $\operatorname{sgn} \sigma = 1, \operatorname{sgn} \tau = -1, \operatorname{sgn} \pi = -1$.

7.67. *a*) $\tau \circ \sigma = 53142$, *b*) $\pi \circ \sigma = 52413$, *c*) $\sigma^{-1} = 32154$, *d*) $\tau^{-1} = 14253$.

7.71. *a*) 30, *b*) 0.

7.72. 17.

7.73. *a*) $-3, -3, -1$; *b*) $-23, -23, -10$.

7.74. *a*) $-2, -17, 73$; *b*) $7, 10, 105$; *c*) $13, 54, 0$.

7.75. $S_1 = -6, S_2 = 13, S_3 = 62, S_4 = -219$.

7.79. *a*) Sí, *b*) no, *c*) sí, *d*) no.

CAPITULO 8

Valores propios y vectores propios. Diagonalización

8.1. INTRODUCCION

Consideremos una matriz n-cuadrada A sobre un cuerpo K. Recordemos (Sección 4.13) que A induce una función f: $K^n \to K^n$ definida según

$$f(X) = AX$$

donde X es cualquier punto (vector columna) en K^n. (Vemos entonces A como la matriz que representa la función f respecto a la base usual E de K^n.)

Supongamos que se elige una nueva base de K^n, digamos

$$S = \{u_1, u_2, \ldots, u_n\}$$

(Geométricamente, S determina un nuevo sistema de coordenadas para K^n.) Sea P la matriz cuyas columnas son los vectores u_1, u_2, ..., u_n. En ese caso (Sección 5.11), P es la matriz de cambio de base desde la usual E hasta la S. Además, en virtud del Teorema 5.27,

$$X' = P^{-1}X$$

proporciona las coordenadas de X en la nueva base S. Asimismo, la matriz

$$B = P^{-1}AP$$

representa la función f en S; esto es, $f(X') = BX'$.

En este capítulo se tratará de contestar a las dos preguntas siguientes:

1. Dada una matriz A, ¿podemos encontrar una matriz no singular P (que represente un nuevo sistema de coordenadas S) de forma que

$$B = P^{-1}AP$$

sea una matriz diagonal? Si la respuesta es afirmativa, diremos que A es *diagonalizable*.

2. Dada una matriz real A, ¿podemos hallar una matriz ortogonal P (que represente un nuevo sistema ortonormal S) de forma que

$$B = P^{-1}AP$$

sea una matriz diagonal? Si la respuesta es afirmativa, diremos que A es *ortogonalmente diagonalizable*.

Recordemos que se dice que dos matrices A y B son similares (ortogonalmente similares) si existe una matriz no singular (ortogonal) P tal que $B = P^{-1}AP$. Lo que está en cuestión es si una matriz dada A es similar (ortogonalmente similar) a una matriz diagonal.

Las respuestas están estrechamente relacionadas con las raíces de ciertos polinomios asociados a A. El cuerpo particular subyacente K juega también un papel importante en esta teoría, puesto que la existencia de raíces de los polinomios depende de K.

8.2. POLINOMIOS DE MATRICES

Consideremos un polinomio $f(t)$ sobre K; por ejemplo,

$$f(t) = a_n t^n + \cdots + a_1 t + a_0$$

Recuérdese que si A es una matriz cuadrada sobre K, definimos

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I$$

donde I es la matriz identidad. En particular decimos que A es una raíz o cero del polinomio $f(t)$ si $f(A) = 0$.

EJEMPLO 8.1. Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y sean $f(t) = 2t^2 - 3t + 7$, $g(t) = t^2 - 5t - 2$. Entonces

$$f(A) = 2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 14 \\ 21 & 39 \end{pmatrix}$$

y

$$g(A) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

De este modo, A es un cero de $g(t)$.

Es aplicable el teorema que ahora enunciamos, demostrado en el Problema 8.26.

Teorema 8.1: Sean f y g polinomios sobre K y A una matriz n-cuadrada sobre K. Se cumple:

i) $(f + g)(A) = f(A) + g(A)$,

ii) $(fg)(A) = f(A)g(A)$,

y, para todo escalar $k \in K$,

iii) $(kf)(A) = kf(A)$.

Más aún, como $f(t)g(t) = g(t)f(t)$ para todo par de polinomios $f(t)$ y $g(t)$,

$$f(A)g(A) = g(A)f(A)$$

Es decir, dos polinomios cualesquiera en la matriz A conmutan.

8.3. POLINOMIO CARACTERISTICO. TEOREMA DE CAYLEY-HAMILTON

Consideremos una matriz n-cuadrada sobre un cuerpo K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

La matriz $tI_n - A$, donde I_n es la matriz identidad n-cuadrada y t un escalar indeterminado, se denomina la *matriz característica* de A:

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix}$$

Su determinante

$$\Delta_A(t) = \det(tI_n - A)$$

que es un polinomio en t, recibe el nombre de *polinomio característico* de A. Asimismo, llamamos a

$$\Delta_A(t) = \det(tI_n - A) = 0$$

la *ecuación característica* de A.

Cada término en el determinante contiene una y sólo una entrada de cada fila y de cada columna, luego el polinomio característico precedente es de la forma

$$\Delta_A(t) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}) + \\ + \text{ términos con a lo sumo } n - 2 \text{ factores de la forma } t - a_{ii}$$

En consecuencia,

$$\Delta_A(t) = t^n - (a_{11} + a_{22} + \cdots + a_{nn})t^{n-1} + \text{ términos de grado menor}$$

Recordemos que la traza de A es la suma de sus elementos diagonales. Siendo así, el polinomio característico $\Delta_A(t) = \det(tI_n - A)$ de A es un polinomio normalizado de grado n y el coeficiente de t^{n-1} es el opuesto de la traza de A. (Un polinomio está *normalizado* si su coeficiente dominante es 1.)

Además, si tomamos $t = 0$ en $\Delta_A(t)$, obtenemos

$$\Delta_A(0) = |-A| = (-1)^n |A|$$

Pero $\Delta_A(0)$ es el término constante del polinomio $\Delta_A(t)$, de modo que el término constante del polinomio característico de la matriz A es $(-1)^n|A|$, siendo n el orden de A.

Establecemos ahora uno de los teoremas más importantes del álgebra lineal (demostrado en el Problema 8.27):

Teorema 8.2 (Cayley-Hamilton): Toda matriz es un cero de su polinomio característico.

EJEMPLO 8.2. Sea $B = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Su polinomio característico es

$$\Delta(t) = |tI - B| = \begin{vmatrix} t-1 & -2 \\ -3 & t-2 \end{vmatrix} = (t-1)(t-2) - 6 = t^2 - 3t - 4$$

Como cabía esperar por el teorema de Cayley-Hamilton, B es un cero de $\Delta(t)$:

$$\Delta(B) = B^2 - 3B - 4I = \begin{pmatrix} 7 & 6 \\ 9 & 10 \end{pmatrix} + \begin{pmatrix} -3 & -6 \\ -9 & -6 \end{pmatrix} + \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Supongamos ahora que A y B son matrices similares, digamos $B = P^{-1}AP$, donde P es invertible. Probemos que A y B tienen el mismo polinomio característico. Utilizando $tI = P^{-1}tIP$,

$$|tI - B| = |tI - P^{-1}AP| = |P^{-1}tIP - P^{-1}AP| =$$
$$= |P^{-1}(tI - A)P| = |P^{-1}||tI - A||P|$$

Dado que los determinantes son escalares y conmutan y que $|P^{-1}||P| = 1$, obtenemos finalmente,

$$|tI - B| = |tI - A|$$

Así hemos demostrado

Teorema 8.3: Las matrices similares tienen el mismo polinomio característico.

POLINOMIOS CARACTERISTICOS DE ORDENES DOS Y TRES

Sea A una matriz de orden dos o tres. En tal caso existe una fórmula sencilla para su polinomio característico $\Delta(t)$. Concretamente:

1. Supongamos $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Entonces

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = t^2 - (\text{tr } A)t + \det(A)$$

(Aquí tr A denota la traza de A, esto es, la suma de sus elementos diagonales.)

2. Supongamos $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Entonces

$$\Delta(t) = t^3 - (a_{11} + a_{22} + a_{33})t^2 + \left(\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right) t -$$

$$- \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = t^3 - (\text{tr } A)t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

(Aquí A_{11}, A_{22}, A_{33} denotan, respectivamente, los cofactores de los elementos diagonales a_{11}, a_{22}, a_{33}.)

Consideremos de nuevo una matriz 3-cuadrada $A = (a_{ij})$. Como se indicó anteriormente,

$$S_1 = \text{tr } A \qquad S_2 = A_{11} + A_{22} + A_{33} \qquad S_3 = \det(A)$$

son los coeficientes de su polinomio característico, con signos alternantes. Por otra parte, cada S_k es la suma de los menores principales de orden k de A. El siguiente teorema, cuya demostración escapa al alcance de este texto, nos dice que este resultado es cierto en general.

Teorema 8.4: Sea A una matriz n-cuadrada. Su polinomio característico es

$$\Delta(t) = t^n - S_1 t^{n-1} + S_2 t^{n-2} - \cdots + (-1)^n S_n$$

donde S_k es la suma de los menores principales de orden k.

POLINOMIO CARACTERISTICO Y MATRICES TRIANGULARES POR BLOQUES

Supongamos que M es una matriz triangular por bloques, por ejemplo, $M = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$, siendo A_1 y A_2 matrices cuadradas. La matriz característica de M,

$$tI - M = \begin{pmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{pmatrix}$$

es también una matriz triangular por bloques, con bloques diagonales $tI - A_1$ y $tI - A_2$. De este modo, según el Teorema 7.12,

$$|tI - M| = \begin{vmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{vmatrix} = |tI - A_1||tI - A_2|$$

Es decir, el polinomio característico de M es el producto de los polinomios característicos de los bloques diagonales A_1 y A_2.

Por inducción puede obtenerse el útil resultado.

Teorema 8.5: Supongamos que M es una matriz triangular por bloques, con bloques diagonales $A_1, A_2, \ldots, A_r$. El polinomio característico de M es el producto de los polinomios característicos de los bloques diagonales A_i, esto es,

$$\Delta_M(t) = \Delta_{A_1}(t)\Delta_{A_2}(t) \cdots \Delta_{A_r}(t)$$

EJEMPLO 8.3. Consideremos la matriz

$$M = \begin{pmatrix} 9 & -1 & 5 & 7 \\ 8 & 3 & 2 & -4 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & -1 & 8 \end{pmatrix}$$

M es una matriz triangular por bloques, con bloques diagonales $A = \begin{pmatrix} 9 & -1 \\ 8 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & 6 \\ -1 & 8 \end{pmatrix}$. Aquí

$$\text{tr } A = 9 + 3 = 12 \qquad \det(A) = 27 + 8 = 35 \qquad \text{y así} \qquad \Delta_A(t) = t^2 - 12t + 35 = (t-5)(t-7)$$

$$\text{tr } B = 3 + 8 = 11 \qquad \det(B) = 24 + 6 = 30 \qquad \text{y así} \qquad \Delta_B(t) = t^2 - 11t + 30 = (t-5)(t-6)$$

De acuerdo con ello, el polinomio característico de M es el producto

$$\Delta_M(t) = \Delta_A(t)\Delta_B(t) = (t-5)^2(t-6)(t-7)$$

8.4. VALORES PROPIOS Y VECTORES PROPIOS

Sea A una matriz n-cuadrada sobre un cuerpo K. Un escalar $\lambda \in K$ se denomina un *valor propio* de A si existe un vector (columna) no nulo $v \in K^n$ para el que

$$Av = \lambda v$$

Todo vector que satisfaga esta relación se llama un *vector propio* de A perteneciente al valor propio λ. Nótese que cada múltiplo escalar kv es a su vez un vector propio, puesto que

$$A(kv) = k(Av) = k(\lambda v) = \lambda(kv)$$

El conjunto E_λ de todos los vectores propios pertenecientes a λ es un subespacio de K^n (Problema 8.16), conocido como *espacio propio* de λ. (Si $\dim E_\lambda = 1$, E_λ recibe el nombre de *recta propia* y λ se llama *factor de escala*.)

Los términos *valor característico* y *vector característico* (o *autovalor* y *autovector*) se utilizan con frecuencia en lugar de valor propio y vector propio.

EJEMPLO 8.4. Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ y sean $v_1 = (2, 3)^T$ y $v_2 = (1, -1)^T$. Entonces

$$Av_1 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 4v_1$$

y

$$Av_2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-1)v_2$$

Así pues, v_1 y v_2 son vectores propios de A pertenecientes, respectivamente, a los valores propios $\lambda_1 = 4$ y $\lambda_2 = -1$ de A.

El teorema enunciado a continuación, demostrado en el Problema 8.28, es la herramienta principal para el cálculo de valores propios y vectores propios (Sección 8.5).

Teorema 8.6: Sea A una matriz n-cuadrada sobre un cuerpo K. Las siguientes afirmaciones son equivalentes:

i) Un escalar $\lambda \in K$ es un valor propio de A.

ii) La matriz $M = \lambda I - A$ es singular.

iii) El escalar λ es una raíz del polinomio característico $\Delta(t)$ de A.

El espacio propio E_λ de λ será el espacio solución del sistema homogéneo $MX = (\lambda I - A)X = 0$.

A veces resulta más conveniente resolver el sistema $(A - \lambda I)X = 0$ que el $(\lambda I - A)X = 0$, cuando se calculan vectores propios. Por supuesto, ambos sistemas conducen al mismo espacio solución.

Algunas matrices pueden no tener valores propios ni, por tanto, vectores propios. No obstante, haciendo uso del teorema fundamental del álgebra (todo polinomio sobre **C** tiene una raíz) y del Teorema 8.6, llegamos al siguiente resultado.

Teorema 8.7. Sea A una matriz n-cuadrada sobre el cuerpo complejo **C**. Entonces A tiene al menos un valor propio.

Supongamos ahora que λ es un valor propio de la matriz A. La *multiplicidad algebraica* de λ es la multiplicidad de λ como raíz del polinomio característico de A. La *multiplicidad geométrica* de λ es la dimensión de su espacio propio.

Es válido el teorema enunciado a continuación, demostrado en el Problema 10.27.

Teorema 8.8: Sea λ un valor propio de una matriz A. La multiplicidad geométrica de λ no excede su multiplicidad algebraica.

MATRICES DIAGONALIZABLES

Se dice que una matriz A es *diagonalizable* (bajo similaridad) si existe una matriz no singular P tal que $D = P^{-1}AP$ es una matriz diagonal, o sea, si A es similar a una matriz diagonal D. El teorema que sigue, demostrado en el Problema 8.29, caracteriza tales matrices.

Teorema 8.9: Una matriz n-cuadrada A es similar a una matriz diagonal D si y sólo si A tiene n vectores propios linealmente independientes. En tal caso, los elementos diagonales de D son los valores propios correspondientes y $D = P^{-1}AP$, siendo P la matriz cuyas columnas son los vectores propios.

Supongamos que una matriz A puede diagonalizarse como antes, digamos $D = P^{-1}AP$ con D diagonal. Entonces A tiene la *factorización diagonal*, extremadamente útil.

$$A = PDP^{-1}$$

Empleando esta factorización, el álgebra de A se reduce a la de la matriz diagonal D, fácilmente manejable. Específicamente, supongamos $D = \text{diag}\,(k_1, k_2, \ldots, k_n)$. En ese caso,

$$A^m = (PDP^{-1})^m = PD^mP^{-1} = P\,\text{diag}\,(k_1^m, \ldots, k_n^m)P^{-1}$$

y, con mayor generalidad, para todo polinomio $f(t)$,

$$f(A) = f(PDP^{-1}) = Pf(D)P^{-1} = P\,\text{diag}\,(f(k_1), \ldots, f(k_n))P^{-1}$$

Más aún, si las entradas diagonales de D son no negativas, la matriz B escrita a continuación es una «raíz cuadrada» de A:

$$B = P\,\text{diag}\,(\sqrt{k_1}, \ldots, \sqrt{k_n})P^{-1}$$

esto es, $B^2 = A$.

EJEMPLO 8.5. Consideremos la matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. De acuerdo con el Ejemplo 8.4, A tiene dos valores propios linealmente independientes $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Tomemos $P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$ y, por tanto, $P^{-1} = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}$. Entonces A es similar a la matriz diagonal

$$B = P^{-1}AP = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

Tal y como cabía esperar, los elementos diagonales 4 y -1 de la matriz diagonal B son los valores propios correspondientes a los vectores propios dados. En particular, A tiene la factorización

$$A = PDP^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix}$$

En consecuencia,

$$A^4 = PD^4P^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 256 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix} = \begin{pmatrix} 103 & 102 \\ 153 & 154 \end{pmatrix}$$

Además, si $f(t) = t^3 - 7t^2 + 9t - 2$,

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} -14 & 0 \\ 0 & -19 \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{2}{5} \end{pmatrix} = \begin{pmatrix} -17 & 2 \\ 3 & -16 \end{pmatrix}$$

Nota: A lo largo de todo este capítulo utilizamos el hecho de que la inversa de la matriz

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{es la matriz} \quad P^{-1} = \begin{pmatrix} d/|P| & -b/|P| \\ -c/|P| & a/|P| \end{pmatrix}$$

Esto es, P^{-1} se obtiene intercambiando los elementos diagonales a y d de P, tomando los opuestos de los elementos no diagonales b y c y dividiendo cada elemento por el determinante $|P|$.

Los dos teoremas siguientes, demostrados en los Problemas 8.30 y 8.31, respectivamente, se utilizarán en lo sucesivo.

Teorema 8.10: Sean $v_1, \ldots, v_n$ vectores propios no nulos de una matriz A, pertenecientes a valores propios distintos $\lambda_1, \ldots, \lambda_n$. Entonces $v_1, \ldots, v_n$ son linealmente independientes.

Teorema 8.11: Supongamos que el polinomio característico de una matriz n-cuadrada A es un producto de n factores distintos, digamos $\Delta(t) = (t - a_1)(t - a_2) \cdots (t - a_n)$. En ese caso, A es similar a una matriz diagonal cuyos elementos diagonales son los a_i.

8.5. CALCULO DE VALORES PROPIOS Y VECTORES PROPIOS. DIAGONALIZACION DE MATRICES

En esta sección se calculan los valores propios y vectores propios de una matriz cuadrada dada A y se determina si existe o no una matriz no singular P tal que $P^{-1}AP$ sea diagonal. En concreto, se aplicará a la matriz A el algoritmo que enseguida exponemos.

Algoritmo 8.5 (Diagonalización)

La entrada es una matriz n-cuadrada A.

Paso 1. Hallar el polinomio característico $\Delta(t)$ de A.

Paso 2. Hallar las raíces de $\Delta(t)$ para obtener los valores propios de A.

Paso 3. Repetir *a*) y *b*) para cada valor propio λ de A:

a) Construir $M = A - \lambda I$ restando λ a los elementos diagonales de A, o $M' = \lambda I - A$ sustituyendo $t = \lambda$ en $tI - A$.

b) Encontrar una base para el espacio solución de $MX = 0$. (Los vectores de la base son vectores propios linealmente independientes de A pertenecientes a λ.)

Paso 4. Considerar la colección $S = \{v_1, v_2, \ldots, v_m\}$ de todos los vectores propios obtenidos en el Paso 3:

a) Si $m \neq n$, A no es diagonalizable.

b) Si $m = n$, sea P la matriz cuyas columnas son los vectores propios $v_1, v_2, \ldots, v_n$. Entonces

$$D = P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ldots & \\ & & & \lambda_n \end{pmatrix}.$$

donde λ_i es el valor propio correspondiente al vector propio v_i.

EJEMPLO 8.6. Apliquemos el algoritmo de diagonalización a $A = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}$.

1. El polinomio característico $\Delta(t)$ de A es el determinante

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-4 & -1 \\ -3 & t+1 \end{vmatrix} = t^2 - 3t - 10 = (t-5)(t+2)$$

Alternativamente, tr $A = 4 - 1 = 3$ y $|A| = -4 - 6 = -10$, de modo que $\Delta(t) = t^2 - 3t - 10$.

2. Hacemos $\Delta(t) = (t-5)(t+2) = 0$. Las raíces $\lambda_1 = 5$ y $\lambda_2 = -2$ son los valores propios de A.

3. i) Hallemos un vector propio v_1 de A perteneciente al valor propio $\lambda_1 = 5$.

Restamos $\lambda_1 = 5$ a los elementos diagonales de A para conseguir la matriz $M = \begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}$.
Los vectores propios pertenecientes a $\lambda_1 = 5$ forman la solución del sistema homogéneo $MX = 0$, es decir,

$$\begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} -x + 2y = 0 \\ 3x - 6y = 0 \end{cases} \quad \text{o} \quad -x + 2y = 0$$

El sistema tiene sólo una solución independiente; por ejemplo, $x = 2$, $y = 1$. Así $v_1 = (2, 1)$ es un vector propio que genera el espacio propio de $\lambda_1 = 5$.

ii) Hallemos un vector propio v_2 de A perteneciente al valor propio $\lambda_2 = -2$.

Restamos -2 (o sumamos 2) a los elementos diagonales de A para conseguir $M = \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix}$ que conduce al sistema homogéneo

$$\begin{cases} 6x + 2y = 0 \\ 3x + y = 0 \end{cases} \quad \text{o} \quad 3x + y = 0$$

El sistema tiene sólo una solución independiente, digamos $x = -1$, $y = 3$. De este modo, $v_2 = (-1, 3)$ es un vector propio que genera el espacio propio de $\lambda_2 = -2$.

4. Sea P la matriz cuyas columnas son los vectores propios precedentes: $P = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$. Entonces $P^{-1} = \begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$ y $D = P^{-1}AP$ es la matriz diagonal que tiene por entradas diagonales los respectivos valores propios:

$$D = P^{-1}AP = \begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$$

De acuerdo con ello, A posee la «factorización diagonal»:

$$A = PDP^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$$

Si $f(t) = t^4 - 4t^3 - 3t^2 + 5$, podemos calcular $f(5) = 55$ y $f(-2) = 41$, luego

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 55 & 0 \\ 0 & 41 \end{pmatrix}\begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix} = \begin{pmatrix} 53 & 4 \\ 6 & 43 \end{pmatrix}$$

EJEMPLO 8.7. Consideremos la matriz $B = \begin{pmatrix} 5 & 1 \\ -4 & 1 \end{pmatrix}$. Aquí tr $B = 5 + 1 = 6$ y $|B| = 5 + 4 = 9$. Por tanto, $\Delta(t) = t^2 - t + 9 = (t - 3)^2$ es el polinomio característico de B. En consecuencia, $\lambda = 3$ es el único valor propio de B.

Restamos $\lambda = 3$ a los elementos diagonales de B obteniendo la matriz $M = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$ que corresponde al sistema homogéneo

$$\begin{cases} 2x + \ \ y = 0 \\ -4x - 2y = 0 \end{cases} \qquad \text{o} \qquad 2x + y = 0$$

El sistema tiene únicamente una solución independiente; por ejemplo, $x = 1$, $y = -2$. Siendo así, $v = (1, -2)$ es el único vector propio de la matriz B. Podemos decir, pues, que B no es diagonalizable, por no existir una base constituida por vectores propios de B.

EJEMPLO 8.8. Consideremos la matriz $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$. Aquí tr $A = 2 - 2 = 0$ y $|A| = -4 + 5 = 1$. Por consiguiente, $\Delta(t) = t^2 + 1$ es el polinomio característico de A. Estudiamos dos casos:

a) A es una matriz sobre el cuerpo real **R**. Entonces $\Delta(t)$ no tiene raíces (reales), de modo que A no tiene valores propios ni vectores propios y por ende no es diagonalizable.

b) A es una matriz sobre el cuerpo complejo **C**. En ese caso, $\Delta(t) = (t - i)(t + i)$ tiene dos raíces, i y $-i$. Así pues, A tiene dos valores propios diferentes, i y $-i$, y dos vectores propios independientes. En consecuencia, existe una matriz no singular P sobre el cuerpo complejo para la cual

$$P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

Por esta razón, A es diagonalizable (sobre **C**).

8.6. DIAGONALIZACION DE MATRICES REALES SIMETRICAS

Hay muchas matrices reales A que no son diagonalizables. De hecho, algunas de ellas pueden no tener ningún valor propio (real). No obstante, si A es una matriz real *simétrica*, estos problemas desaparecen. A saber:

Teorema 8.12: Sea A una matriz real simétrica. Toda raíz λ de su polinomio característico es real.

Teorema 8.13: Sea A una matriz real simétrica. Supongamos que u y v son vectores propios no nulos de A pertenecientes a valores propios distintos λ_1 y λ_2. Entonces u y v son ortogonales, o sea, $\langle u, v \rangle = 0$.

Los dos teoremas anteriores nos llevan al resultado fundamental:

Teorema 8.14: Sea A una matriz real simétrica. Existe una matriz ortogonal P tal que $D = P^{-1}AP$ es diagonal.

Podemos elegir las columnas de la matriz P precedente como vectores propios ortogonales normalizados de A, en cuyo caso las entradas diagonales de D serán los valores propios correspondientes.

EJEMPLO 8.9. Sea $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. Hallemos una matriz ortogonal P tal que $P^{-1}AP$ sea diagonal. Aquí tr $A = 2 + 5 = 7$ y $|A| = 10 - 4 = 6$. Por tanto, $\Delta(t) = t^2 - 7t + 6 = (t-6)(t-1)$ es el polinomio de A. Los valores propios de A son 6 y 1. Restando $\lambda = 6$ de la diagonal de A obtenemos el sistema de ecuaciones lineales homogéneo asociado

$$-4x - 2y = 0 \qquad -2x - y = 0$$

Una solución no nula es $v_1 = (1, -2)$. Acto seguido restamos $\lambda = 1$ de la diagonal de la matriz A para llegar al sistema homogéneo asociado

$$x - 2y = 0 \qquad -2x + 4y = 0$$

Una solución no nula es $v_2 = (2, 1)$. Como podía esperarse por el Teorema 8.13, v_1 y v_2 son ortogonales. Los normalizamos para conseguir los vectores ortonormales

$$\mathbf{u}_1 = (1/\sqrt{5}, -2/\sqrt{5}) \qquad \mathbf{u}_2 = (2/\sqrt{5}, 1/\sqrt{5})$$

Finalmente, sea P la matriz cuyas columnas son u_1 y u_2, respectivamente. Entonces

$$\mathbf{P} = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \qquad \text{y} \qquad \mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

Como era de esperar, las entradas diagonales de $P^{-1}AP$ son los valores propios correspondientes a las columnas de P.

APLICACION A LAS FORMAS CUADRATICAS

Recordemos (Sección 4.12) que una forma cuadrática real $q(x_1, x_2, \ldots, x_n)$ puede expresarse en la forma matricial

$$q(X) = X^T AX$$

donde $X = (x_1, \ldots, x_n)^T$ y A es una matriz real simétrica. Recordemos, asimismo, que bajo un cambio de variables $X = PY$, con $Y = (y_1, \ldots, y_n)$ y P una matriz no singular, la forma cuadrática adopta la forma

$$q(Y) = Y^T BY$$

siendo $B = P^T AP$. (Así B es congruente a A.)

Ahora bien, si P es una matriz ortogonal, $P^T = P^{-1}$. En tal caso, $B = P^T AP = P^{-1}AP$, con lo que B es ortogonalmente similar a A. En consecuencia, el método precedente para diagonalizar una matriz real simétrica A puede emplearse para diagonalizar una forma cuadrática q bajo un cambio de coordenadas ortogonal, como se indica a continuación.

Algoritmo 8.6 (Diagonalización ortogonal)

La entrada es una forma cuadrática $q(X)$.

Paso 1. Hallar la matriz simétrica A que represente q y calcular su polinomio característico $\Delta(t)$.

Paso 2. Hallar los valores propios de A, que son las raíces de $\Delta(t)$.

Paso 3. Encontrar una base ortogonal del espacio propio de cada uno de los valores propios λ del Paso 2.

Paso 4. Normalizar todos los vectores propios del Paso 3, que pasarán así a formar una base ortonormal de $\mathbf{R}^n$.

Paso 5. Sea P la matriz cuyas columnas son los vectores propios normalizados del Paso 4.

Entonces $X = PY$ es el cambio de coordenadas ortogonal requerido. Las entradas diagonales de P^TAP serán los valores propios $\lambda_1, \ldots, \lambda_n$ asociados a las columnas de P.

8.7. POLINOMIO MINIMO

Sea A una matriz n-cuadrada sobre un cuerpo K y denotemos por $J(A)$ la colección de todos los polinomios $f(t)$ para los cuales $f(A) = 0$. [Nótese que $J(A)$ no es vacío, ya que el polinomio característico $\Delta(t)$ de A pertenece a $J(A)$.] Sea $m(t)$ el polinomio normalizado de grado mínimo en $J(A)$. Entonces $m(t)$ recibe el nombre de *polinomio mínimo* de A. [Tal polinomio $m(t)$ existe y es único (Problema 8.25).]

Teorema 8.15: El polinomio mínimo $m(t)$ de A es divisor de todo polinomio que tenga A como cero. En particular, $m(t)$ es divisor del polinomio característico $\Delta(t)$ de A.

(La demostración se da en el Problema 8.32.) Existe una relación más fuerte incluso entre $m(t)$ y $\Delta(t)$.

Teorema 8.16: Los polinomios característico y mínimo de una matriz A tienen los mismos factores irreducibles.

Este teorema, demostrado en el Problema 8.33 *b*), no dice que $m(t) = \Delta(t)$, sino, únicamente, que todo factor irreducible de uno debe ser divisor del otro. En particular, como todo factor lineal es irreducible, $m(t)$ y $\Delta(t)$ tienen los mismos factores lineales y por ende las mismas raíces. Así tenemos:

Teorema 8.17: Un escalar λ es un valor propio de una matriz A si y sólo si es una raíz de su polinomio mínimo.

EJEMPLO 8.10. Hallemos el polinomio mínimo $m(t)$ de $A = \begin{pmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{pmatrix}$.

Primero calculamos el polinomio característico $\Delta(t)$ de A:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -2 & 5 \\ -3 & t-7 & 15 \\ -1 & -2 & t+4 \end{vmatrix} = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$$

Alternativamente, $\Delta(t) = t^3 - (\text{tr } A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 5t^2 + 7t - 3 = (t-1)^2(t-3)$ (donde A_{ii} es el cofactor de a_{ii} en A).

El polinomio mínimo $m(t)$ debe ser divisor de $\Delta(t)$. Además, todo factor irreducible de $\Delta(t)$, esto es, $t - 1$ y $t - 3$, debe ser a su vez un factor de $m(t)$. De este modo, $m(t)$ es exactamente uno de los siguientes polinomios:

$$f(t) = (t-3)(t-1) \qquad \text{o} \qquad g(t) = (t-3)(t-1)^2$$

Sabemos, por el Teorema de Cayley-Hamilton, que $g(A) = \Delta(A) = 0$, luego sólo necesitamos comprobar $f(t)$. Tenemos

$$f(A) = (A-I)(A-3I) = \begin{pmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{pmatrix}\begin{pmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Así pues, $f(t) = m(t) = (t-1)(t-3) = t^2 - 4t + 3)$ es el polinomio mínimo de A.

EJEMPLO 8.11. Consideremos la matriz n-cuadrada, con $a \neq 0$:

$$M = \begin{pmatrix} \lambda & a & 0 & \dots & 0 & 0 \\ 0 & \lambda & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & a \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Nótese que los elementos de M son λ en la diagonal, a en la superdiagonal y 0 en cualquier otra posición. Esta matriz es importante en álgebra lineal, especialmente cuando $a = 1$. Puede probarse que

$$f(t) = (t - \lambda)^n$$

es tanto el polinomio característico como el polinomio mínimo de M.

EJEMPLO 8.12. Consideremos un polinomio normalizado arbitrario $f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. Sea A la matriz n-cuadrada con 1 en la subdiagonal, los opuestos de los coeficientes en la última columna y 0 en cualquier otro lugar:

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Se dice que A es la *matriz acompañante* del polinomio $f(t)$. El polinomio mínimo $m(t)$ y el polinomio característico $\Delta(t)$ de esta matriz acompañante A son ambos iguales a $f(t)$.

POLINOMIO MINIMO Y MATRICES DIAGONALES POR BLOQUES

Puede utilizarse el siguiente teorema, demostrado en el Problema 8.34.

Teorema 8.18. Supongamos que M es una matriz diagonal por bloques, con bloques diagonales $A_1, A_2, \ldots, A_r$. El polinomio mínimo de M es igual al mínimo común múltiplo (MCM) de los polinomios mínimos de los bloques diagonales A_i.

Nota: Subrayamos que este teorema se aplica a matrices diagonales por bloques, mientras que el Teorema 8.5, su análogo para polinomios característicos, se aplica a matrices triangulares por bloques.

EJEMPLO 8.13. Hallemos el polinomio característico $\Delta(t)$ y el polinomio mínimo $m(t)$ de la matriz

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

Nótese que A es una matriz diagonal por bloques, con bloques diagonales

$$A_1 = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix} \qquad A_3 = (7)$$

Por esta razón, $\Delta(t)$ es el producto de los polinomios característicos $\Delta_1(t)$, $\Delta_2(t)$ y $\Delta_3(t)$ de A_1, A_2 y A_3, respectivamente. Siendo triangulares A_1 y A_3, $\Delta_1(t) = (t-2)^2$ e $\Delta_3(t) = (t-7)$. Además,

$$\Delta_2(t) = t^2 - (\text{tr } A_2)t + |A_2| = t^2 - 9t + 14 = (t-2)(t-7)$$

Por tanto, $\Delta(t) = (t-2)^3(t-7)^2$. [Como cabía esperar, grado $\Delta(t) = 5$.]

Los polinomios mínimos $m_1(t)$, $m_2(t)$ y $m_3(t)$ de los bloques diagonales A_1, A_2 y A_3, respectivamente, coinciden con los polinomios característicos; esto es,

$$m_1(t) = (t-2)^2 \qquad m_2(t) = (t-2)(t-7) \qquad m_3(t) = t-7$$

Pero $m(t)$ es igual al mínimo común múltiplo de $m_1(t)$, $m_2(t)$, $m_3(t)$. De este modo, $m(t) = (t-2)^2(t-7)$.

PROBLEMAS RESUELTOS

POLINOMIOS DE MATRICES. POLINOMIO CARACTERISTICO

8.1. Sea $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$. Hallar $f(A)$, donde: *a)* $f(t) = t^2 - 3t + 7$, *b)* $f(t) = t^2 - 6t + 13$.

a) $f(A) = A^2 - 3A + 7I = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$

$= \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix} + \begin{pmatrix} -3 & 6 \\ -12 & -15 \end{pmatrix} + \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 12 & 9 \end{pmatrix}$

b) $f(A) = A^2 - 6A + 13I = \begin{pmatrix} -7 & -12 \\ 24 & 17 \end{pmatrix} + \begin{pmatrix} -6 & 12 \\ -24 & -30 \end{pmatrix} + \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

[Así A es una raíz de $f(t)$.]

8.2. Encontrar el polinomio característico $\Delta(t)$ de la matriz $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$.

Construyamos la matriz característica $tI - A$:

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} + \begin{pmatrix} -2 & 3 \\ -5 & -1 \end{pmatrix} = \begin{pmatrix} t-2 & 3 \\ -5 & t-1 \end{pmatrix}$$

El polinomio característico $\Delta(t)$ de A es el determinante:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 3 \\ -5 & t-1 \end{vmatrix} = (t-2)(t-1) + 15 = t^2 - 3t + 17$$

Alternativamente, $\operatorname{tr} A = 2 + 1 = 3$ y $|A| = 2 + 15 = 17$, luego $\Delta(t) = t^2 - 3t + 17$.

8.3. Hallar el polinomio característico $\Delta(t)$ de la matriz $A = \begin{pmatrix} 1 & 6 & -2 \\ -3 & 2 & 0 \\ 0 & 3 & -4 \end{pmatrix}$.

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -6 & 2 \\ 3 & t-2 & 0 \\ 0 & -3 & t+4 \end{vmatrix} = (t-1)(t-2)(t+4) - 18 + 18(t+4) = t^3 + t^2 - 8t + 62$$

Alternativamente, $\operatorname{tr} A = 1 + 2 - 4 = -1$, $A_{11} = \begin{vmatrix} 2 & 0 \\ 3 & -4 \end{vmatrix} = -8$, $A_{22} = \begin{vmatrix} 1 & -2 \\ 0 & -4 \end{vmatrix} = -4$, $A_{33} = \begin{vmatrix} 1 & 6 \\ -3 & 2 \end{vmatrix} =$ $= 2 + 18 = 20$, $A_{11} + A_{22} + A_{33} = -8 - 4 + 20 = 8$ y $|A| = -8 + 18 - 72 = -62$. Así pues,

$$\Delta(t) = t^3 - (\operatorname{tr} A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 + t^2 - 8t + 62$$

8.4. Hallar el polinomio característico $\Delta(t)$ de cada una de las siguientes matrices:

a) $R = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 8 & -6 \\ 0 & 0 & 3 & -5 \\ 0 & 0 & 0 & 4 \end{pmatrix}$, b) $S = \begin{pmatrix} 2 & 5 & 7 & -9 \\ 1 & 4 & -6 & 4 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 2 & 3 \end{pmatrix}$

a) Como R es triangular, $\Delta(t) = (t-1)(t-2)(t-3)(t-4)$.

b) Nótese que S es triangular por bloques, con bloques diagonales $A_1 = \begin{pmatrix} 2 & 5 \\ 1 & 4 \end{pmatrix}$ y $A_2 = \begin{pmatrix} 6 & -5 \\ 2 & 3 \end{pmatrix}$. Por tanto,

$$\Delta(t) = \Delta_{A_1}(t)\Delta_{A_2}(t) = (t^2 - 6t + 3)(t^2 - 9t + 28)$$

VALORES PROPIOS Y VECTORES PROPIOS

8.5. Sea $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. Encontrar: *a*) todos los valores propios de A y los espacios propios correspondientes, *b*) una matriz invertible P tal que $D = P^{-1}AP$ sea diagonal, *c*) A^5 y $f(A)$ para $f(t) = t^4 - 3t^3 - 7t^2 + 6t - 15$.

a) Formemos la matriz característica $tI - A$ de A:

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} t-1 & -4 \\ -2 & t-3 \end{pmatrix}$$

El polinomio característico $\Delta(t)$ de A es su determinante:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -4 \\ -2 & t-3 \end{vmatrix} = t^2 - 4t - 5 = (t-5)(t+1)$$

Alternativamente, tr $A = 1 + 3 = 4$ y $|A| = 3 - 8 = -5$, luego $\Delta(t) = t^2 - 4t - 5$. Las raíces $\lambda_1 = 5$ y $\lambda_2 = -1$ del polinomio característico $\Delta(t)$ son los valores propios de A.

Obtengamos los vectores propios de A pertenecientes al valor propio $\lambda_1 = 5$. Para ello sustituimos $t = 5$ en la matriz característica [1], llegando a la matriz $M = \begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}$. Los vectores propios pertenecientes a $\lambda_1 = 5$ constituyen la solución del sistema homogéneo $MX = 0$, o sea,

$$\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} 4x - 4y = 0 \\ -2x + 2y = 0 \end{cases} \quad \text{o} \quad x - y = 0$$

El sistema sólo tiene una solución independiente; por ejemplo, $x = 1$, $y = 1$. De este modo, $v_1 = (1, 1)$ es un vector propio que genera el espacio propio de $\lambda_1 = 5$.

Obtengamos los vectores propios de A pertenecientes al valor propio $\lambda_2 = -1$. Sustituimos $t = -1$ en $tI - A$, llegando a la matriz $M = \begin{pmatrix} -2 & -4 \\ -2 & -4 \end{pmatrix}$ que proporciona el sistema homogéneo

$$\begin{cases} -2x - 4y = 0 \\ -2x - 4y = 0 \end{cases} \quad \text{o} \quad x + 2y = 0$$

El sistema sólo tiene una solución independiente; por ejemplo, $x = 2$, $y = -1$. Así $v_2 = (2, -1)$ es un vector propio que genera el espacio propio de $\lambda_2 = -1$.

b) Sea P la matriz cuyas columnas son los vectores propios anteriores: $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. Entonces $D = P^{-1}AP$ es la matriz diagonal cuyas entradas diagonales son los respectivos valores propios:

$$D = P^{-1}AP = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

[*Nota*: Aquí P es la matriz de cambio de base desde la usual E de $\mathbf{R}^2$ hasta la base $S = (v_1, v_2)$. Por tanto, B es la representación matricial de la función determinada por A en esta nueva base.]

c) Usemos la factorización diagonal de A,

$$A = PDP^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

para obtener ($5^5 = 3125$ y $(-1)^5 = 1$):

$$A^5 = PD^5P^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3125 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1041 & 2084 \\ 1042 & 2083 \end{pmatrix}$$

Asimismo, como $f(5) = 90$ y $f(-1) = -24$:

$$f(A) = Pf(D)P^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 90 & 0 \\ 0 & -24 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 14 & 76 \\ 38 & 52 \end{pmatrix}$$

8.6. Hallar todos los valores propios y un conjunto máximo de vectores propios linealmente independientes para las matrices:

$$a)\quad A = \begin{pmatrix} 5 & 6 \\ 3 & -2 \end{pmatrix} \qquad b)\quad C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}$$

¿Cuáles de ellas pueden diagonalizarse? En su caso, determinar la matriz no singular P requerida para ello.

a) Calculamos el polinomio característico $\Delta(t) = t^2 - 3t - 28 = (t-7)(t+4)$. Los valores propios de A son, pues, $\lambda_1 = 7$ y $\lambda_2 = -4$.

i) Restamos $\lambda_1 = 7$ a los elementos diagonales de A obteniendo $M = \begin{pmatrix} -2 & 6 \\ 3 & -9 \end{pmatrix}$, correspondiente al sistema

$$\begin{cases} -2x + 6y = 0 \\ 3x - 9y = 0 \end{cases} \qquad \text{o} \qquad x - 3y = 0$$

Aquí $v_1 = (3, 1)$ es una solución no nula (que genera el espacio solución) y por ende es el vector propio de $\lambda_1 = 7$.

ii) Restamos $\lambda_2 = -4$ a los elementos diagonales de A consiguiendo $M = \begin{pmatrix} 9 & 6 \\ 3 & 2 \end{pmatrix}$, que corresponde al sistema $3x + 2y = 0$. Aquí $v_2 = (2, -3)$ es una solución y por consiguiente un vector propio de $\lambda_2 = -4$.

Entonces $S = \{v_1 = (3, 1), v_2 = (2, -3)\}$ es un conjunto máximo de vectores propios linealmente independientes de A. Dado que S es una base de $\mathbf{R}^2$, A es diagonalizable. Sea P la matriz cuyas columnas son v_1 y v_2. En tal caso,

$$P = \begin{pmatrix} 3 & 2 \\ 1 & -3 \end{pmatrix} \qquad \text{y} \qquad P^{-1}AP = \begin{pmatrix} 7 & 0 \\ 0 & -4 \end{pmatrix}$$

b) Hallamos $\Delta(t) = t^2 - 8t + 16 = (t-4)^2$. El único valor propio es, así, $\lambda = 4$. Restamos $\lambda = 4$ a los elementos diagonales de C para obtener $M = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$, correspondiente al sistema homogé-

neo $x - y = 0$. Una solución no nula del sistema es $v = (1, 1)$, por lo que v es un vector propio de C perteneciente a $\lambda = 4$. Al no haber otros valores propios, el conjunto de un solo elemento $S = \{v = (1, 1)\}$ es un conjunto máximo de vectores propios linealmente independientes. Además, C no es diagonalizable, porque el número de vectores propios linealmente independientes no es igual a la dimensión del espacio vectorial $\mathbf{R}^2$. En particular, no existe la matriz no singular P citada.

8.7. Sea $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$. Encontrar: *a*) todos los valores propios de A y los vectores propios asociados; *b*) una matriz no singular P tal que $D = P^{-1}AP$ sea diagonal; *c*) A^6; *d*) una «raíz cuadrada» positiva de A, es decir, una matriz B con autovalores no negativos tal que $B^2 = A$.

a) Aquí $\Delta(t) = t^2 - \text{tr } A + |A| = t^2 - 5t + 4 = (t-1)(t-4)$. Por consiguiente, $\lambda_1 = 1$ y $\lambda_2 = 4$ son valores propios de A. Hallemos los vectores propios correspondientes:

i) Restamos $\lambda_1 = 1$ a los elementos diagonales de A llegando a $M = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, que nos lleva al sistema homogéneo $x + 2y = 0$. Como $v_1 = (2, -1)$ es una solución no nula del sistema, será un vector propio de A perteneciente a $\lambda_1 = 1$.

ii) Restamos $\lambda_2 = 4$ a los elementos diagonales de A obteniendo $M = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$, que corresponde al sistema homogéneo $x - y = 0$. Aquí $v_2 = (1, 1)$ es una solución no nula y por tanto un valor propio de A perteneciente a $\lambda_2 = 4$.

b) Sea P la matriz cuyas columnas son v_1 y v_2. Entonces

$$P = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{y} \quad D = P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

c) Utilicemos la factorización diagonal de A,

$$A = PDP^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

para obtener

$$A^6 = PD^6P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4096 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1366 & 2730 \\ 1365 & 2731 \end{pmatrix}$$

d) Aquí $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 2 \end{pmatrix}$ son raíces cuadradas de D, luego

$$B = P\sqrt{D}\,P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{5}{3} \end{pmatrix}$$

es la raíz cuadrada positiva de A (similar a una matriz diagonal con elementos no negativos).

8.8. Supóngase $A = \begin{pmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix}$. Calcular: *a*) el polinomio característico $\Delta(t)$ de A, *b*) los valores propios de A, *c*) un conjunto máximo de vectores propios linealmente indepen-

dientes de A. *d*) ¿Es diagonalizable A? En caso afirmativo, encontrar P tal que $P^{-1}AP$ sea diagonal.

a) Tenemos

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-4 & -1 & 1 \\ -2 & t-5 & 2 \\ -1 & -1 & t-2 \end{vmatrix} = t^3 - 11t^2 + 39t - 45$$

Alternativamente, $\Delta(t) = t^3 - (\operatorname{tr} A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 11t^2 + 39t - 45$. (Aquí A_{ii} es el cofactor de a_{ii} en la matriz A.)

b) Suponiendo que $\Delta(t)$ tenga una raíz racional, ésta debe estar entre ± 1, ± 3, ± 5, ± 9, ± 15, ± 45. Probando conseguimos

$$\begin{array}{r|l} 3 & 1 - 11 + 39 - 45 \\ & \quad\;\; 3 - 24 + 45 \\ \hline & 1 - \;\; 8 + 15 + \;\; 0 \end{array}$$

Así $t = 3$ es una raíz de $\Delta(t)$ y $t - 3$ un factor que da

$$\Delta(t) = (t-3)(t^2 - 8t + 15) = (t-3)(t-5)(t-3) = (t-3)^2(t-5)$$

En consecuencia, $\lambda_1 = 3$ y $\lambda_2 = 5$ son los valores propios de A.

c) Hallemos vectores propios linealmente independientes para cada valor propio.

i) Restamos $\lambda_1 = 3$ a los elementos diagonales de A llegando a $M = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{pmatrix}$, correspondiente al sistema homogéneo $x + y - z = 0$. Aquí $u = (1, -1, 0)$ y $v = (1, 0, 1)$ son dos soluciones independientes.

ii) Restamos $\lambda_2 = 5$ a los elementos diagonales para obtener $M = \begin{pmatrix} -1 & 1 & -1 \\ 2 & 0 & -2 \\ 1 & 1 & -3 \end{pmatrix}$, que lleva asociado el sistema homogéneo

$$\begin{cases} -x + y - z = 0 \\ 2x - 2z = 0 \\ x + y - 3z = 0 \end{cases} \quad \text{o} \quad \begin{cases} x - z = 0 \\ y - 2z = 0 \end{cases}$$

La única variable libre es z. Una solución es $w = (1, 2, 1)$.

De este modo, $\{u = (1, -1, 0),\ v = (1, 0, 1),\ w = (1, 2, 1)\}$ es un conjunto máximo de vectores propios linealmente independientes de A.

Nota: Se han elegido los vectores u y v de forma que fueran soluciones independientes del sistema homogéneo $x + y - z = 0$. Por otra parte, w es automáticamente independiente de u y v, por pertenecer a un valor propio diferente de A. Los tres vectores son, pues, linealmente independientes.

d) A es diagonalizable, ya que tiene tres vectores propios linealmente independientes. Sea P la matriz con columnas u, v, w. En ese caso,

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{y} \quad P^{-1}AP = \begin{pmatrix} 3 & & \\ & 3 & \\ & & 5 \end{pmatrix}$$

8.9. Supóngase $B = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$. Hallar: *a*) el polinomio característico $\Delta(t)$ y los valores propios de B, *b*) un conjunto máximo de vectores propios linealmente independientes de B. (*c*) ¿Es diagonalizable B? En caso afirmativo, encontrar P tal que $P^{-1}BP$ sea diagonal.

a) Tenemos

$$\Delta(t) = |tI - B| = \begin{vmatrix} t+3 & -1 & 1 \\ 7 & t-5 & 1 \\ 6 & -6 & t+2 \end{vmatrix} = t^3 - 12t - 16$$

Por consiguiente, $\Delta(t) = (t+2)^2(t-4)$. Así $\lambda_1 = -2$ y $\lambda_2 = 4$ son los valores propios de B.

b) Hallemos una base para el espacio propio de cada valor propio.

i) Sustituimos $t = -2$ en $tI - B$ obteniendo el sistema homogéneo

$$\begin{pmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} x - y + z = 0 \\ 7x - 7y + z = 0 \\ 6x - 6y \quad = 0 \end{cases} \quad \text{o} \quad \begin{cases} x - y + z = 0 \\ x - y \quad = 0 \end{cases}$$

El sistema tiene únicamente una solución linealmente independiente, por ejemplo, $x = 1$, $y = 1$, $z = 0$. Siendo así, $u = (1, 1, 0)$ forma una base del espacio propio de $\lambda_1 = -2$.

ii) Sustituimos $t = 4$ en $tI - B$ para llegar al sistema homogéneo

$$\begin{pmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} 7x - y + z = 0 \\ 7x - y + z = 0 \\ 6x - 6y + 6z = 0 \end{cases} \quad \text{o} \quad \begin{cases} 7x - y + z = 0 \\ x = 0 \end{cases}$$

El sistema tiene sólo una solución linealmente independiente, digamos $x = 0$, $y = 1$, $z = 1$. Por tanto, $v = (0, 1, 1)$ constituye una base del espacio propio de $\lambda_2 = 4$.

De este modo, $S = \{u, v\}$ es un conjunto máximo de vectores propios linealmente independientes de B.

c) Dado que B tiene un máximo de dos vectores propios linealmente independientes, no es similar a una matriz diagonal, o sea, B no es diagonalizable.

8.10. Hallar las multiplicidades algebraica y geométrica del valor propio $\lambda_1 = -2$ de la matriz B del Problema 8.9.

La multiplicidad algebraica de λ_1 es dos. Sin embargo, la multiplicidad geométrica de λ_1 es uno, puesto que $\dim E_{\lambda_1} = 1$.

8.11. Sea $A = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$. Encontrar todos los valores propios y los correspondientes vectores propios de A suponiendo que A es una matriz real. ¿Es diagonalizable A? En caso afirmativo, hallar P tal que $P^{-1}AP$ sea diagonal.

El polinomio característico de A es $\Delta(t) = t^2 + 1$, que no tiene raíces en **R**. Así A, vista como una matriz real, no tiene valores propios ni vectores propios y por ende A no es diagonalizable sobre **R**.

8.12. Repetir el Problema 8.11 suponiendo ahora que A es una matriz sobre el cuerpo complejo **C**.

El polinomio característico de A sigue siendo $\Delta(t) = t^2 + 1$. (No depende del cuerpo K.) Sobre **C**, $\Delta(t)$ es factorizable; concretamente, $\Delta(t) = t^2 + 1 = (t - i)(t + i)$. Por tanto, $\lambda_1 = i$ y $\lambda_2 = -i$ son valores propios de A.

i) Sustituimos $t = i$ en $tI - A$ para obtener el sistema homogéneo

$$\begin{pmatrix} i-1 & 1 \\ -2 & i+1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} (i-1)x + y = 0 \\ -2x + (i+1)y = 0 \end{cases} \quad \text{o} \quad (i-1)x + y = 0$$

El sistema tiene sólo una solución linealmente independiente, por ejemplo, $x = 1$, $y = 1 - i$. De este modo, $v_1 = (1, 1 - i)$ es un vector propio que genera el espacio propio de $\lambda_1 = i$.

ii) Sustituimos $t = -i$ en $tI - A$ para obtener el sistema homogéneo

$$\begin{pmatrix} -i-1 & 1 \\ -2 & -i-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} (-i-1)x + y = 0 \\ -2x + (-i-1)y = 0 \end{cases} \quad \text{o} \quad (-i-1)x + y = 0$$

El sistema tiene sólo una solución linealmente independiente, por ejemplo, $x = 1$, $y = 1 + i$. Así $v_2 = (1, 1 + i)$ es un vector propio que genera el espacio propio de $\lambda_2 = -i$.

Como matriz compleja, A es diagonalizable. Sea P la matriz cuyas columnas son v_1 y v_2. Entonces

$$P = \begin{pmatrix} 1 & 1 \\ 1-i & 1+i \end{pmatrix} \quad \text{y} \quad P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

8.13. Sea $B = \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}$. Hallar: *a*) todos los valores propios de B y los vectores propios asociados; *b*) una matriz invertible P tal que $D = P^{-1}BP$ sea diagonal; *c*) B^6.

a) Aquí $\Delta(t) = t^2 - \operatorname{tr} B + |B| = t^2 - 3t - 10 = (t - 5)(t + 2)$. Por tanto, $\lambda_1 = 5$ y $\lambda_2 = -2$ son los valores propios de B.

i) Restamos $\lambda_1 = 5$ a todos los elementos diagonales de B llegando a $M = \begin{pmatrix} -3 & 4 \\ 3 & -4 \end{pmatrix}$, que corresponde al sistema homogéneo $3x - 4y = 0$. Una solución no nula es $v_1 = (4, 3)^T$.

ii) Restamos $\lambda_2 = -2$ (o sumamos 2) a los elementos diagonales de B para conseguir $M = \begin{pmatrix} 4 & 4 \\ 3 & 3 \end{pmatrix}$, que corresponde al sistema homogéneo $x + y = 0$ que tiene una solución no nula $v_2 = (1, -1)^T$.

(Como B tiene dos vectores propios independientes, es diagonalizable.)

b) Sea P la matriz cuyas columnas son v_1 y v_2. En tal caso,

$$P = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \quad \text{y} \quad D = P^{-1}BP = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$$

c) Empleamos la factorización diagonal de B,

$$B = PDP^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix}$$

para obtener ($5^6 = 15625$, $(-2)^6 = 64$):

$$B^6 = PD^6P^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 15{,}625 & 0 \\ 0 & 64 \end{pmatrix}\begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix} = \begin{pmatrix} 8956 & 8892 \\ 6669 & 6733 \end{pmatrix}$$

8.14. Determinar si A es o no diagonalizable, donde $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$.

Siendo A triangular, los valores propios de A son sus elementos diagonales 1, 2 y 3. Dado que son distintos, A tiene tres vectores propios independientes, de modo que es similar a una matriz diagonal (Teorema 8.11). (Señalamos que aquí no hay necesidad de calcular vectores propios para poder afirmar que A es diagonalizable. Si quisiéramos hallar P tal que $P^{-1}AP$ fuera diagonal, tendríamos que calcular vectores propios.)

8.15. Supóngase que A y B son matrices n-cuadradas.

a) Probar que 0 es un valor propio de A si y sólo si A es singular.

b) Probar que AB y BA tienen los mismos valores propios.

c) Supóngase que A es no singular (invertible) y que λ es un valor propio de A. Probar que λ^{-1} es un valor propio de A^{-1}.

d) Probar que A y su traspuesta A^T tienen el mismo polinomio característico.

a) Tenemos que 0 es un valor propio de A si y sólo si existe un vector no nulo v tal que $A(v) = 0v = 0$, esto es, si y sólo si A es singular.

b) Según la parte *a)* y el hecho de que el producto de matrices no singulares es no singular, las siguientes afirmaciones son equivalentes: i) 0 es un valor propio de AB, ii) AB es singular, iii) A o B es singular, iv) BA es singular, v) 0 es un valor propio de BA.

Supongamos ahora que λ es un valor propio distinto de cero de AB. Entonces existe un vector no nulo v tal que $ABv = \lambda v$. Tomemos $w = Bv$. Siendo $\lambda \neq 0$ y $v \neq 0$,

$$Aw = ABv = \lambda v \neq 0 \qquad \text{y así} \qquad w \neq 0$$

Pero w es un vector propio de BA perteneciente al valor propio λ, ya que

$$BAw = BABv = B\lambda v = \lambda Bv = \lambda w$$

Por consiguiente, λ es un valor propio de BA. De forma similar, todo valor propio no nulo de BA es también un valor propio de AB.

Así pues, AB y BA tienen los mismos valores propios.

c) De acuerdo con la parte *a)*, $\lambda \neq 0$. Por definición de valor propio, existe un vector no nulo v para el que $A(v) = \lambda v$. Aplicando A^{-1} a ambos miembros obtenemos $v = A^{-1}(\lambda v) = \lambda A^{-1}(v)$. Por tanto, $A^{-1}(v) = \lambda^{-1}v$; es decir, λ^{-1} es un valor propio de A^{-1}.

d) Como una matriz y su traspuesta tienen el mismo determinante, $|tI - A| = |(tI - A)^T| = |tI - A^T|$. De este modo, A y A^T tienen el mismo polinomio característico.

8.16. Sea λ un valor propio de una matriz n-cuadrada A sobre K. Sea E_λ el espacio propio de λ, o sea, el conjunto de todos los vectores propios de A pertenecientes a λ. Mostrar que E_λ es un subespacio de K^n, esto es, mostrar que: *a)* si $v \in E_\lambda$, necesariamente $kv \in E_\lambda$ para todo escalar $k \in K$, *b)* si u, $v \in E_\lambda$, necesariamente $u + v \in E_\lambda$.

a) Como $v \in E_\lambda$, tenemos $A(v) = \lambda v$. Entonces

$$A(kv) = kA(v) = k(\lambda v) = \lambda(kv)$$

Por tanto, $kv \in E_2$. [Hemos de admitir el vector nulo de K^n como «vector propio» asociado a $k = 0$ para hacer de E_λ un subespacio.]

b) Como $u, v \in E_\lambda$, tenemos $A(u) = \lambda v$ y $A(v) = \lambda v$. Entonces

$$A(u + v) = A(u) + A(v) = \lambda u + \lambda v = \lambda(u + v)$$

Por tanto, $u + v \in E_\lambda$.

DIAGONALIZACION DE MATRICES REALES SIMETRICAS Y DE FORMAS CUADRATICAS REALES

8.17. Sea $A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$. Encontrar una matriz (real) ortogonal P para la cual $P^T AP$ sea diagonal.

El polinomio característico $\Delta(t)$ de A es

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-3 & -2 \\ -2 & t-3 \end{vmatrix} = t^2 - 6t + 5 = (t-5)(t-1)$$

y así los valores propios de A son 5 y 1.

Restamos $\lambda = 5$ de la diagonal de la matriz A para conseguir el sistema de ecuaciones lineales homogéneo correspondiente

$$-2x + 2y = 0 \qquad 2x - 2y = 0$$

Una solución no nula es $v_1 = (1, 1)$. Normalizamos v_1 para hallar la solución unitaria $u_1 = (1/\sqrt{2}, 1/\sqrt{2})$.

A continuación restamos $t = 1$ de la diagonal de la matriz A obteniendo el sistema de ecuaciones lineales homogéneo

$$2x + 2y = 0 \qquad 2x + 2y = 0$$

Una solución no nula es $v_2 = (1, -1)$. Normalizamos v_2 para hallar la solución unitaria $u_2 = (1/\sqrt{2}, -1/\sqrt{2})$.

Finalmente, sea P la matriz cuyas columnas son u_1 y u_2, respectivamente; entonces

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \qquad \text{y} \qquad P^T AP = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

Tal y como cabía esperar, las entradas diagonales de $P^T AP$ son los valores propios de A.

8.18. Supóngase $C = \begin{pmatrix} 11 & -8 & 4 \\ -8 & -1 & -2 \\ 4 & -2 & -4 \end{pmatrix}$. Hallar: *a*) el polinomio característico $\Delta(t)$ de C; *b*) los valores propios de C o, en otras palabras, las raíces de $\Delta(t)$; *c*) un conjunto máximo S de vectores propios ortogonales no nulos de C; *d*) una matriz ortogonal P tal que $P^{-1}AP$ sea diagonal.

a) Tenemos

$$\Delta(t) = t^3 - (\text{tr } C)t^2 + (C_{11} + C_{22} + C_{33})t - |C| = t^3 - 6t^2 - 135t - 400$$

[Aquí C_{ii} es el cofactor de c_{ii} en $C = (c_{ij})$.]

b) Si $\Delta(t)$ tiene una raíz racional, debe ser divisor de 400. Probando con $t = -5$,

$$\begin{array}{r|rrrr} -5 & 1 - & 6 - & 135 - & 400 \\ & - & 5 + & 55 + & 400 \\ \hline & 1 - & 11 - & 80 + & 0 \end{array}$$

Por tanto, $t + 5$ es un factor de $\Delta(t)$ y

$$\Delta(t) = (t + 5)(t^2 - 11t - 80) = (t + 5)^2(t - 16)$$

En consecuencia, los valores propios de C son $\lambda = -5$ (con multiplicidad dos) y $\lambda = 16$ (con multiplicidad uno).

c) Hallemos una base ortogonal para cada espacio propio.

Restamos $\lambda = -5$ a todos los elementos diagonales de C para obtener el sistema homogéneo

$$16x - 8y + 4z = 0 \qquad -8x + 4y - 2z = 0 \qquad 4x - 2y + z = 0$$

Esto es, $4x - 2y + z = 0$. El sistema tiene dos soluciones independientes. Una de ellas es $v_1 = (0, 1, 2)$. Buscamos una segunda solución $v_2 = (a, b, c)$ que sea ortogonal a v_1, es decir, tal que

$$4a - 2b + c = 0 \qquad \text{y además} \qquad b - 2c = 0$$

Una de estas soluciones es $v_2 = (-5, -8, 4)^T$.

Restamos $\lambda = 16$ a los elementos diagonales de C llegando al sistema homogéneo.

$$-5x - 8y + 4z = 0 \qquad -8x - 17y - 2z = 0 \qquad 4x - 2y - 20z = 0$$

Este sistema proporciona una solución no nula $v_3 = (4, -2, 1)$. (Como podía esperarse por el Teorema 8.13, el vector propio v_3 es ortogonal a v_1 y v_2.)

Entonces v_1, v_2, v_3 constituyen un conjunto máximo de vectores propios ortogonales no nulos de C.

d) Normalizamos v_1, v_2, v_3 obteniendo la base ortonormal

$$u_1 = (0, 1/\sqrt{5}, 2/\sqrt{5}) \qquad u_2 = (-5/\sqrt{105}, -8/\sqrt{105}, 4/\sqrt{105}) \qquad u_3 = (4/\sqrt{21}, -2/\sqrt{21}, 1/\sqrt{21})$$

Sabemos que P es la matriz cuyas columnas son u_1, u_2, u_2. Por tanto,

$$P = \begin{pmatrix} 0 & -5/\sqrt{105} & 4/\sqrt{21} \\ 1/\sqrt{5} & -8/\sqrt{105} & -2/\sqrt{21} \\ 2/\sqrt{5} & 4/\sqrt{105} & 1/\sqrt{21} \end{pmatrix} \qquad \text{y} \qquad P^TCP = \begin{pmatrix} -5 & & \\ & -5 & \\ & & 16 \end{pmatrix}$$

8.19. Sea $q(x, y) = 3x^2 - 6xy + 11y^2$. Encontrar un cambio de coordenadas ortogonal que diagonalice q.

Hallamos la matriz simétrica que representa q y su polinomio característico $\Delta(t)$:

$$A = \begin{pmatrix} 3 & -3 \\ -3 & 11 \end{pmatrix} \qquad \text{y} \qquad \Delta(t) = \begin{vmatrix} t-3 & 3 \\ 3 & t-11 \end{vmatrix} = t^2 - 14t + 24 = (t-2)(t-12)$$

Los valores propios son 2 y 12, luego una forma diagonal de q es

$$q(x', y') = 2x'^2 + 12y'^2$$

El cambio de coordenadas asociado se consigue hallando un conjunto correspondiente de vectores propios de A.

Restamos $t = 2$ de la diagonal de la matriz A para obtener el sistema homogéneo

$$-9x - 3y = 0 \qquad -3x - y = 0$$

Una solución no nula es $v_1 = (3, 1)$. Seguidamente sustituimos $t = 12$ en la matriz $tI - A$ y llegamos al sistema

$$9x + 3y = 0 \qquad 3x + y = 0$$

Una solución no nula es $v_2 = (-1, 3)$. Normalizamos v_1 y v_2 para conseguir la base ortonormal

$$u_1 = (3/\sqrt{10}, 1/\sqrt{10}) \qquad u_2 = (-1/\sqrt{10}, 3/\sqrt{10})$$

La matriz de cambio de base y el cambio de coordenadas pedido son:

$$P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{o} \quad \begin{cases} x = (3x' - y')/\sqrt{10} \\ y = (x' + 3y')/\sqrt{10} \end{cases}$$

También pueden expresarse x' e y' en términos de x e y, haciendo uso de $P^{-1} = P^T$, esto es,

$$x' = (3x + y)/\sqrt{10} \qquad y' = (-x + 3y)/\sqrt{10}$$

8.20. Considérese la forma cuadrática $q(x, y, z) = 3x^2 + 2xy + 3y^2 + 2xz + 2yz + 3z^2$. Hallar:

a) La matriz simétrica A que representa q y su polinomio característico $\Delta(t)$.

b) Los valores propios de A o, si se prefiere, las raíces de $\Delta(t)$.

c) Un conjunto máximo de vectores propios de A ortogonales no nulos.

d) Un cambio de coordenadas ortogonal que diagonalice q.

a) Recordemos que $A = (a_{ij})$ es la matriz simétrica en la que a_{ii} es el coeficiente de x_i^2 y $a_{ij} = a_{ji}$ es la mitad del coeficiente de $x_i x_j$. Así pues,

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{y} \quad \Delta(t) = \begin{vmatrix} t-3 & -1 & -1 \\ -1 & t-3 & -1 \\ -1 & -1 & t-3 \end{vmatrix} = t^3 - 9t^2 + 24t - 20$$

b) Si $\Delta(t)$ tiene una raíz racional, debe ser divisor de la constante 20 o, dicho de otro modo, debe estar entre ± 1, ± 2, ± 4, ± 5, ± 10, ± 20. Probando con $t = 2$,

$$\begin{array}{r|r} 2 & 1 - 9 + 24 - 20 \\ & 2 - 14 + 20 \\ \hline & 1 - 7 + 10 + \ \ 0 \end{array}$$

Así $t - 2$ es un factor de $\Delta(t)$, con

$$\Delta(t) = (t - 2)(t^2 - 7t + 10) = (t - 2)^2(t - 5)$$

Por consiguiente, los valores propios de A son 2 (con multiplicidad dos) y 5 (con multiplicidad uno).

c) Hallemos una base ortogonal de cada espacio propio.

Restamos $\lambda = 2$ a los elementos diagonales de A obteniendo el sistema homogéneo correspondiente

$$x + y + z = 0 \qquad x + y + z = 0 \qquad x + y + z = 0$$

Esto es, $x + y + z = 0$. El sistema tiene dos soluciones independientes. Una de tales soluciones es $v_1 = (0, 1, -1)$. Buscamos una segunda solución $v_2 = (a, b, c)$ que sea ortogonal a v_1, o sea, tal que

$$a + b + c = 0 \qquad \text{y además} \qquad b - c = 0$$

Por ejemplo, $v_2 = (2, -1, -1)$. De este modo, $v_1 = (0, 1, -1)$, $v_2 = (2, -1, -1)$ forman una base ortogonal para el espacio propio de $\lambda = 2$.

Restamos $\lambda = 5$ a los elementos diagonales de A llegando al sistema homogéneo

$$-2x + y + z = 0 \qquad x - 2y + z = 0 \qquad x + y - 2z = 0$$

Este sistema proporciona una solución no nula $v_3 = (1, 1, 1)$. (Como podía esperarse por el Teorema 8.13, el vector propio v_3 es ortogonal a v_1 y v_2.)

Entonces v_1, v_2, v_3 integran un conjunto máximo de vectores propios ortogonales no nulos de A.

d) Normalizamos v_1, v_2, v_3 consiguiendo la base ortonormal

$$u_1 = (0, 1/\sqrt{2}, -1/\sqrt{2}) \qquad u_2 = (2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6}) \qquad u_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

Sea P la matriz cuyas columnas son u_1, u_2, u_3. En tal caso

$$P = \begin{pmatrix} 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \qquad \text{y} \qquad P^T A P = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 5 \end{pmatrix}$$

de donde el cambio de coordenadas requerido es

$$\begin{aligned} x &= \frac{2y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}} \\ y &= \frac{x'}{\sqrt{2}} - \frac{y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}} \\ z &= -\frac{x'}{\sqrt{2}} - \frac{y'}{\sqrt{6}} + \frac{z'}{\sqrt{3}} \end{aligned}$$

Bajo este cambio de coordenadas, q adopta la forma diagonal

$$q(x', y', z') = 2x'^2 + 2y'^2 + 5z'^2$$

POLINOMIO MINIMO

8.21. Encontrar el polinomio mínimo $m(t)$ de la matriz $A = \begin{pmatrix} 2 & -1 & 1 \\ 6 & -3 & 4 \\ 3 & -2 & 3 \end{pmatrix}$.

Comenzamos calculando el polinomio característico $\Delta(t)$ de A:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 1 & -1 \\ -6 & t+3 & -4 \\ -3 & 2 & t-3 \end{vmatrix} = t^3 - 4t^2 + 5t - 2 = (t-2)(t-1)^2$$

Alternativamente, $\Delta(t) = t^3 - (\operatorname{tr} A)t^2 + (A_{11} + A_{22} + A_{33})t - |A| = t^3 - 4t^2 + 5t - 2 = (t-2)(t-1)^2$. (Aquí A_{ii} es el cofactor de a_{ii} en A.)

El polinomio mínimo $m(t)$ debe ser divisor de $\Delta(t)$. Asimismo, todo factor irreducible de $\Delta(t)$, es decir, $t-2$ y $t-1$, debe serlo también de $m(t)$. De este modo, $m(t)$ es exactamente uno de los siguientes polinomios:

$$f(t) = (t-2)(t-1) \quad \text{o} \quad g(t) = (t-2)(t-1)^2$$

Sabemos, por el Teorema de Cayley-Hamilton, que $g(A) = \Delta(A) = 0$, luego sólo necesitamos comprobar $f(t)$. Tenemos

$$f(A) = (A - 2I)(A - I) = \begin{pmatrix} 2 & -2 & 2 \\ 6 & -5 & 4 \\ 3 & -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & -2 & 2 \\ 6 & -4 & 4 \\ 3 & -2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Así $f(t) = m(t) = (t-2)(t-1) = t^2 - 3t + 2$ es el polinomio mínimo de A.

8.22. Hallar el polinomio mínimo $m(t)$ de la matriz (donde $a \neq 0$): $B = \begin{pmatrix} \lambda & a & 0 \\ 0 & \lambda & a \\ 0 & 0 & \lambda \end{pmatrix}$.

El polinomio característico de B es $\Delta(t) = (t - \lambda)^3$. [Adviértase que $m(t)$ es exactamente uno de los polinomios $t - \lambda$, $(t-\lambda)^2$ o $(t-\lambda)^3$.] Encontramos que $(B - \lambda I)^2 \neq 0$, luego $m(t) = \Delta(t) = (t-\lambda)^3$.

(*Nota:* Esta matriz es un caso particular del Ejemplo 8.11 y el Problema 8.61.)

8.23. Encontrar el polinomio mínimo $m(t)$ de la siguiente matriz: $M' = \begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$.

Aquí M' es diagonal por bloques con bloques diagonales

$$A' = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{y} \quad B' = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$$

El polinomio característico y mínimo de A' es $f(t) = (t-4)^3$, y el de B', $g(t) = (t-4)^2$. Siendo así, $\Delta(t) = f(t)g(t) = (t-4)^5$ es el polinomio característico de M', pero $m(t) = \text{MCM}\,[f(t), g(t)] = (t-4)^3$ (donde 3 es el tamaño del bloque mayor) es el polinomio mínimo de M'.

8.24. Determinar una matriz A cuyo polinomio mínimo sea:

a) $f(t) = t^3 - 8t^2 + 5t + 7$, *b*) $f(t) = t^4 - 3t^3 - 4t^2 + 5t + 6$

Sea A la matriz acompañante (véase el Ejemplo 8.12) de $f(t)$. En tal caso,

$$a)\quad A = \begin{pmatrix} 0 & 0 & -7 \\ 1 & 0 & -5 \\ 0 & 1 & 8 \end{pmatrix}, \qquad b)\quad A = \begin{pmatrix} 0 & 0 & 0 & -6 \\ 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

(*Nota:* El polinomio $f(t)$ es también el polinomio característico de A.)

8.25. Probar que el polinomio mínimo de una matriz A existe y es único.

De acuerdo con el Teorema de Cayley-Hamilton, A es un cero de algún polinomio no nulo (véase también el Problema 8.37). Sea n el mínimo grado para el que existe un polinomio $f(t)$ con $f(A) = 0$. Dividiendo $f(t)$ por su coeficiente dominante obtenemos un polinomio normalizado $m(t)$ de grado n que tiene A como cero. Supongamos que $m'(t)$ es otro polinomio normalizado de grado n para el cual $m'(A) = 0$. Entonces la diferencia $m(t) - m'(t)$ es un polinomio no nulo de grado menor que n que tiene A como cero. Esto contradice la suposición inicial referente a n, luego $m(t)$ es el único polinomio mínimo.

DEMOSTRACION DE LOS TEOREMAS

8.26. Demostrar el Teorema 8.1.

Supongamos $f = a_n t^n + \cdots + a_1 t + a_0$ y $g = b_m t^m + \cdots + b_1 t + b_0$. Entonces, por definición,

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I \qquad \text{y} \qquad g(A) = b_m A^m + \cdots + b_1 A + b_0 I$$

i) Supongamos $m \leqslant n$ y sea $b_i = 0$ si $i > m$. En ese caso,

$$f + g = (a_n + b_n)t^n + \cdots + (a_1 + b_1)t + (a_0 + b_0)$$

luego

$$\begin{aligned}(f + g)(A) &= (a_n + b_n)A^n + \cdots + (a_1 + b_1)A + (a_0 + b_0)I = \\ &= a_n A^n + b_n A^n + \cdots + a_1 A + b_1 A + a_0 I + b_0 I = f(A) + g(A)\end{aligned}$$

ii) Por definición, $fg = c_{n+m}t^{n+m} + \cdots + c_1 t + c_0 = \sum_{k=0}^{n+m} c_k t^k$, donde

$$c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}$$

De aquí $(fg)(A) = \sum_{k=0}^{n+m} c_k A^k$ y

$$f(A)g(A) = \left(\sum_{i=0}^{n} a_i A^i\right)\left(\sum_{j=0}^{m} b_j A^j\right) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j A^{i+j} = \sum_{k=0}^{n+m} c_k A^k = (fg)(A)$$

iii) Por definición, $kf = ka_n t^n + \cdots + ka_1 t + ka_0$, y así

$$(kf)(A) = ka_n A^n + \cdots + ka_1 A + ka_0 I = k(a_n A^n + \cdots + a_1 A + a_0 I) = kf(A)$$

iv) Por ii), $g(A)f(A) = (gf)(A) = (fg)(A) = f(A)g(A)$.

8.27. Demostrar el Teorema 8.2 (Cayley-Hamilton).

Sean A una matriz n-cuadrada arbitraria y $\Delta(t)$ su polinomio característico; por ejemplo,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Ahora denotemos por $B(t)$ el adjunto clásico de la matriz $tI - A$. Los elementos de $B(t)$ son cofactores de la matriz $tI - A$ y por tanto son polinomios en t de grado no superior a $n - 1$. De este modo,

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

donde las B_i son matrices n-cuadradas sobre K, independientes de t. Según la propiedad fundamental del adjunto clásico (Teorema 7.9)

$$(tI - A)B(t) = |tI - A|I$$

o
$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Eliminando los paréntesis e igualando los coeficientes de las mismas potencias de t,

$$\begin{aligned} B_{n-1} &= I \\ B_{n-2} - AB_{n-1} &= a_{n-1}I \\ B_{n-3} - AB_{n-2} &= a_{n-2}I \\ &\cdots\cdots\cdots \\ B_0 - AB_1 &= a_1 I \\ -AB_0 &= a_0 I \end{aligned}$$

Multiplicando las ecuaciones matriciales precedentes por $A^n, A^{n-1}, \ldots, A, I$, respectivamente,

$$\begin{aligned} A^n B_{n-1} &= A^n \\ A^{n-1}B_{n-2} - A^n B_{n-1} &= a_{n-1}A^{n-1} \\ A^{n-2}B_{n-3} - A^{n-1}B_{n-2} &= a_{n-2}A^{n-2} \\ &\cdots\cdots\cdots \\ AB_0 - A^2 B_1 &= a_1 A \\ -AB_0 &= a_0 I \end{aligned}$$

Sumando las ecuaciones matriciales anteriores,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

En otras palabras, $\Delta(A) = 0$, que es el Teorema de Cayley-Hamilton.

8.28. Demostrar el Teorema 8.6.

El escalar λ es un valor propio de A si y sólo si existe un vector no nulo v tal que

$$Av = \lambda v \qquad \text{o} \qquad (\lambda I)v - Av = 0 \qquad \text{o} \qquad (\lambda I - A)v = 0$$

o $M = \lambda I - A$ es singular. En tal caso, λ es una raíz de $\Delta(t) = |tI - A|$. Asimismo, v está en el espacio propio E_λ de λ si y sólo si se verifican las relaciones anteriores, luego v es una solución de $(\lambda I - A)X = 0$.

8.29. Demostrar el Teorema 8.9.

Supongamos que A tiene n vectores propios linealmente independientes $v_1, v_2, \ldots, v_n$ con valores propios correspondientes $\lambda_1, \lambda_2, \ldots, \lambda_n$. Sea P la matriz cuyas columnas son $v_1, \ldots, v_n$. Entonces P es no singular. Las columnas de AP son $Av_1, \ldots, Av_n$. Pero $Av_k = \lambda_k v_k$. Por consiguiente, las columnas de AP son $\lambda_1 v_1, \ldots, \lambda_n v_n$. Por otra parte, sea $D = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$, es decir, la matriz diagonal con entradas diagonales λ_k. En ese caso, PD es también una matriz con columnas $\lambda_k v_k$. En consecuencia,

$$AP = PD \qquad \text{y por tanto} \qquad D = P^{-1}AP$$

como se pedía.

Recíprocamente, supongamos que existe una matriz no singular P para la cual

$$P^{-1}AP = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n) = D \qquad \text{y así} \qquad AP = PD$$

Sean $v_1, v_2, \ldots, v_n$ los vectores columna de P. Entonces las columnas de AP son Av_k y las de PD $\lambda_k v_k$. De acuerdo con esto, dado que $AP = PD$, tenemos

$$Av_1 = \lambda_1 v_1,\ Av_2 = \lambda_2 v_2, \ldots, Av_n = \lambda_n v_n$$

Además, al ser P no singular, $v_1, v_2, \ldots, v_n$ son no nulos y por tanto son vectores propios de A pertenecientes a los valores propios que son los elementos diagonales de D. Más aún, son linealmente independientes. Queda así demostrado el teorema.

8.30. Demostrar el Teorema 8.10.

Se demuestra por inducción en n. Si $n = 1$, v_1 es necesariamente independiente porque $v_1 \neq 0$. Supongamos $n > 1$ y

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \qquad [1]$$

donde los a_i son escalares. Multiplicando [1] por A,

$$a_1 Av_1 + a_2 Av_2 + \cdots + a_n Av_n = A0 = 0$$

Por hipótesis, $Av_i = \lambda_i v_i$. De este modo, al sustituir, obtenemos

$$a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_n \lambda_n v_n = 0 \qquad [2]$$

Por otra parte, multiplicando [1] por λ_n llegamos a

$$a_1 \lambda_n v_1 + a_2 \lambda_n v_2 + \cdots + a_n \lambda_n v_n = 0 \qquad [3]$$

Restar [3] a [2] proporciona

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0$$

Por inducción, $v_1, v_2, \ldots, v_{n-1}$ son linealmente independientes, luego cada uno de los coeficientes precedentes es 0. Como los λ_i son distintos, $\lambda_i - \lambda_n \neq 0$ para $i \neq n$. De aquí $a_1 = \cdots = a_{n-1} = 0$. Sustituyendo esto en [1] conseguimos $a_n v_n = 0$ y por ende $a_n = 0$. Los v_i son, pues, linealmente independientes.

8.31. Demostrar el Teorema 8.11.

Según el Teorema 8.6, los a_i son valores propios de A. Sean v_i los correspondientes vectores propios. De acuerdo con el Teorema 8.10, los v_i son linealmente independientes, luego forman una base de K^n. Siendo así, A es diagonalizable por el Teorema 8.9.

8.32. Demostrar el Teorema 8.15.

Supongamos que $f(t)$ es un polinomio para el que $f(A) = 0$. Por el algoritmo de la división sabemos que existen polinomios $q(t)$ y $r(t)$ para los cuales $f(t) = m(t)q(t) + r(t)$ con $r(t) = 0$ o grado $r(t) <$ grado $m(t)$. Sustituyendo $t = A$ en esta ecuación y utilizando $f(A) = 0$ y $m(A) = 0$ obtenemos $r(A) = 0$. Si $r(t) \neq 0$, debe ser un polinomio de grado menor que el de $m(t)$ que tiene A como cero. Esto contradice la definición de polinomio mínimo. Así pues, $r(t) = 0$ y $f(t) = m(t)q(t)$, o sea, $m(t)$ es divisor de $f(t)$.

8.33. Sea $m(t)$ el polinomio mínimo de una matriz n-cuadrada A.

a) Probar que el polinomio característico de A es divisor de $(m(t))^n$.

b) Demostrar el Teorema 8.16.

a) Supongamos $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Consideremos las siguientes matrices:

$$\begin{aligned} B_0 &= I \\ B_1 &= A + c_1 I \\ B_2 &= A^2 + c_1 A + c_2 I \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I \end{aligned}$$

Entonces

$$\begin{aligned} B_0 &= I \\ B_1 - AB_0 &= c_1 I \\ B_2 - AB_1 &= c_2 I \\ &\cdots\cdots\cdots\cdots \\ B_{r-1} - AB_{r-2} &= c_{r-1} I \end{aligned}$$

Además,

$$\begin{aligned} -AB_{r-1} &= c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) = \\ &= c_r I - m(A) = \\ &= c_r I \end{aligned}$$

Tomemos

$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

En ese caso,

$$\begin{aligned} (tI - A) \cdot B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) = \\ &= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} = \\ &= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} tI + c_r I = \\ &= m(t) I \end{aligned}$$

Tomar el determinante de ambos miembros nos conduce a $|tI - A||B(t)| = |m(t)I| = (m(t))^n$. Como $|B(t)|$ es un polinomio, $|tI - A|$ es divisor de $(m(t))^n$; esto es, el polinomio característico de A es divisor de $(m(t))^n$.

b) Supongamos que $f(t)$ es un polinomio irreducible. Si $f(t)$ es divisor de $m(t)$, al ser $m(t)$ divisor de $\Delta(t)$, $f(t)$ debe serlo también. Por otra parte, si $f(t)$ es divisor de $\Delta(t)$, según *a*), $f(t)$ es divisor de $(m(t))^n$. Pero $f(t)$ es irreducible, de modo que $f(t)$ debe ser también divisor de $m(t)$. Por tanto, $m(t)$ y $\Delta(t)$ tienen los mismos factores irreducibles.

8.34. Demostrar el Teorema 8.18.

Demostrémoslo para el caso $r = 2$. El teorema general se deriva fácilmente por inducción. Supongamos $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, donde A y B son matrices cuadradas. Debemos probar que el polinomio mínimo $m(t)$ de M es el mínimo común múltiplo de los polinomios mínimos $g(t)$ y $h(t)$ de A y B, respectivamente.

Como $m(t)$ es el polinomio mínimo de M, $m(M) = \begin{pmatrix} m(A) & 0 \\ 0 & m(B) \end{pmatrix} = 0$ y por consiguiente $m(A) = 0$ y $m(B) = 0$. Dado que $g(t)$ es el polinomio mínimo de A, $g(t)$ es divisor de $m(t)$. De forma similar, $h(t)$ es divisor de $m(t)$. Así $m(t)$ es múltiplo de $g(t)$ y $h(t)$.

Ahora sea $f(t)$ otro múltiplo de $g(t)$ y $h(t)$; entonces $f(M) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$. Pero $m(t)$ es el polinomio mínimo de M, luego es divisor de $f(t)$. De este modo, $m(t)$ es el mínimo común múltiplo de $g(t)$ y $h(t)$.

8.35. Supóngase que A es una matriz real simétrica vista como matriz sobre $\mathbf{C}$.

a) Demostrar que $\langle Au, v \rangle = \langle u, Av \rangle$ para el producto interno en $\mathbf{C}^n$.

b) Demostrar los Teoremas 8.12 y 8.13 para la matriz A.

a) Usamos el hecho de que el producto interno en $\mathbf{C}^n$ está definido por $\langle u, v \rangle = u^T \bar{v}$. Puesto que A es real simétrica, $A = A^T = \bar{A}$. Así pues,

$$\langle Au, v \rangle = (Au)^T \bar{v} = u^T A^T \bar{v} = u^T \bar{A} \bar{v} = u^T \overline{Av} = \langle u, Av \rangle$$

b) Utilizamos el hecho de que en $\mathbf{C}^n$ $\langle ku, v \rangle = k\langle u, v \rangle$ pero $\langle u, kv \rangle = \bar{k}\langle u, v \rangle$.

1. Existe $v \neq 0$ tal que $Av = \lambda v$. En ese caso,

$$\lambda\langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda}\langle v, v \rangle$$

Pero $\langle v, v \rangle \neq 0$ porque $v \neq 0$. De este modo, $\lambda = \bar{\lambda}$ y por tanto A es real.

2. Aquí $Au = \lambda_1 u$ y $Av = \lambda_2 v$. Por 1, λ_2 es real. Entonces

$$\lambda_1\langle u, v \rangle = \langle \lambda_1 u, v \rangle = \langle Au, v \rangle = \langle u, Av \rangle = \langle u, \lambda_2 v \rangle = \bar{\lambda}_2\langle u, v \rangle = \lambda_2\langle u, v \rangle$$

Como $\lambda_1 \neq \lambda_2$, tenemos $\langle u, v \rangle = 0$.

PROBLEMAS VARIOS

8.36. Supóngase que A es una matriz simétrica 2×2 con valores propios 1 y 9 y que $u = (1, 3)^T$ es un vector propio perteneciente al valor propio 1. Hallar: *a*) un vector propio v perteneciente al valor propio 9, *b*) la matriz A, *c*) una raíz cuadrada de A, es decir, una matriz B tal que $B^2 = A$.

a) Siendo A simétrica, v debe ser ortogonal a u. Tomemos $v = (-3, 1)^T$.

b) Sea P la matriz cuyas columnas son los vectores propios u y v. Tenemos, por la factorización diagonal de A,

$$A = PDP^{-1} = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}\begin{pmatrix} \frac{1}{10} & \frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{pmatrix} = \begin{pmatrix} \frac{41}{5} & -\frac{12}{5} \\ -\frac{12}{5} & \frac{9}{5} \end{pmatrix}$$

(Alternativamente, A es la matriz para la cual $Au = u$ y $Av = 9v$.)

c) Empleamos la factorización diagonal de A para obtener

$$B = P\sqrt{D}\,P^{-1} = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} \frac{1}{10} & \frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{pmatrix} = \begin{pmatrix} \frac{14}{5} & -\frac{3}{5} \\ -\frac{3}{5} & \frac{6}{5} \end{pmatrix}$$

8.37. Sea A una matriz n-cuadrada. Sin usar el teorema de Cayley-Hamilton, mostrar que A es una raíz de algún polinomio no nulo.

Sea $N = n^2$. Consideremos las $N + 1$ matrices siguientes:

$$I, A, A^2, \ldots, A^N$$

Recuérdese que el espacio vectorial V de las matrices $n \times n$ tiene dimensión $N = n^2$. Las $N + 1$ matrices precedentes son, pues, linealmente dependientes. Siendo así, existen escalares $a_0, a_1, a_2, \ldots, a_N$, no todos nulos, para los que

$$a_N A^N + \cdots + a_1 A + a_0 I = 0$$

De este modo, A es una raíz del polinomio $f(t) = a_N t^N + \cdots + a_1 t + a_0$.

8.38. Supóngase que A es una matriz n-cuadrada. Probar las siguientes afirmaciones:

a) A es no singular si y sólo si el término constante en el polinomio mínimo de A es distinto de cero.

b) Si A es no singular, A^{-1} es igual a un polinomio en A de grado no superior a n.

a) Supongamos $f(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$ es el polinomio mínimo (característico) de A. Entonces son equivalentes: i) A es no singular, ii) 0 no es una raíz de $f(t)$, iii) el término constante a_0 no es cero. Así pues, la afirmación es cierta.

b) Sea $m(t)$ el polinomio mínimo de A. En ese caso, $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$, donde $r \leqslant n$. Como A es no singular, $a_0 \neq 0$, según la parte *a*). Tenemos

$$m(A) = A^r + a_{r-1}A^{r-1} + \cdots + a_1 A + a_0 I = 0$$

De este modo,

$$-\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1 I)A = I$$

En consecuencia,

$$A^{-1} = -\frac{1}{a_0}(A^{r-1} + a_{r-1}A^{r-2} + \cdots + a_1 I)$$

8.39. Sean F una extensión de un cuerpo K y A una matriz n-cuadrada sobre K. Nótese que A puede verse también como una matriz $\hat{A}$ sobre F. Claramente, $|tI - A| = |tI - \hat{A}|$, esto

es, A y $\hat{A}$ tienen el mismo polinomio característico. Mostrar que A y $\hat{A}$ tienen además el mismo polinomio mínimo.

Sean $m(t)$ y $m'(t)$ los polinomios mínimos de A y $\hat{A}$, respectivamente. Ahora $m'(t)$ es divisor de todo polinomio sobre F que tenga A como cero. Dado que $m(t)$ tiene A como cero y puede verse como un polinomio sobre F, $m'(t)$ es divisor de $m(t)$. Probemos ahora que $m(t)$ es divisor de $m'(t)$.

Siendo $m'(t)$ un polinomio sobre F, que es una extensión de K, podemos escribir

$$m'(t) = f_1(t)b_1 + f_2(t)b_2 + \cdots + f_n(t)b_n$$

donde los $f_i(t)$ son polinomios sobre K y $b_1, \ldots, b_n$ pertenecen a F y son linealmente independientes sobre K. Tenemos

$$m'(A) = f_1(A)b_1 + f_2(A)b_2 + \cdots + f_n(A)b_n = 0 \qquad [1]$$

Denotemos por $a_{ij}^{(k)}$ la entrada ij de $f_k(A)$. La ecuación matricial anterior implica que, para todo par (i, j),

$$a_{ij}^{(1)}b_1 + a_{ij}^{(2)}b_2 + \cdots + a_{ij}^{(n)}b_n = 0$$

Como los b_i son linealmente independientes sobre K y los $a_{ij}^{(k)} \in K$, todo $a_{ij}^{(k)} = 0$. Entonces

$$f_1(A) = 0, \; f_2(A) = 0, \; \ldots, \; f_n(A) = 0$$

Al ser los $f_i(t)$ polinomios sobre K que tienen A como cero y puesto que $m(t)$ es el polinomio mínimo de A como matriz sobre K, $m(t)$ es divisor de cada uno de los $f_i(t)$. En consecuencia, por [1], $m(t)$ debe ser divisor de $m'(t)$. Pero dos polinomios normalizados divisores uno del otro son necesariamente iguales. O sea, $m(t) = m'(t)$ como se requería.

PROBLEMAS SUPLEMENTARIOS

POLINOMIOS DE MATRICES

8.40. Sean $f(t) = 2t^2 - 5t + 6$ y $g(t) = t^3 - 2t^2 + t + 3$. Hallar $f(A)$, $g(A)$, $f(B)$ y $g(B)$, donde

$$A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

8.41. Sea $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Hallar A^2, A^3, A^n.

8.42. Sea $\boldsymbol{B} = \begin{pmatrix} 8 & 12 & 0 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{pmatrix}$. Hallar una matriz real A tal que $B = A^3$.

8.43. Probar que, para cualquier matriz n-cuadrada A, $(P^{-1}AP)^n = P^{-1}A^nP$, donde P es invertible. Con mayor generalidad, probar que $f(P^{-1}AP) = P^{-1}f(A)P$ para todo polinomio $f(t)$.

8.44. Sea $f(t)$ un polinomio arbitrario. Mostrar que *a*) $f(A^T) = (f(A))^T$, y *b*) si A es simétrica, entonces $f(A)$ es simétrica.

VALORES PROPIOS Y VECTORES PROPIOS

8.45. Sea $A = \begin{pmatrix} 5 & 6 \\ -2 & -2 \end{pmatrix}$. Hallar: *a*) todos los valores propios y vectores propios linealmente independientes; *b*) P tal que $D = P^{-1}AP$ sea diagonal; *c*) A^{10} y $f(A)$, donde $f(t) = t^4 - 5t^3 + 7t^2 - 2t + 5$; *d*) B tal que $B^2 = A$.

8.46. Para cada una de las siguientes matrices, encontrar todos los valores propios y una base cada espacio propio:

$$a)\ A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}, \qquad b)\ B = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}, \qquad c)\ C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cuando sea posible, hallar matrices invertibles P_1, P_2 y P_3 tales que $P_1^{-1}AP_1$, $P_2^{-1}BP_2$ y $P_3^{-1}CP_3$ sean diagonales.

8.47. Considérense las matrices $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}$. Encontrar todos los valores propios y vectores propios linealmente independientes suponiendo que: *a*) A y B son matrices sobre el cuerpo real **R**, *b*) A y B son matrices sobre el cuerpo complejo **C**.

8.48. Supóngase que v es un vector propio no nulo de dos matrices A y B. Mostrar que v es también vector propio de la matriz $kA + k'B$, donde k y k' son escalares cualesquiera.

8.49. Supóngase que v es un vector propio no nulo de una matriz A perteneciente al valor propio λ. Probar que para $n > 0$ v es también un vector propio de A^n perteneciente a λ^n.

8.50. Supóngase que λ es un valor propio de una matriz A. Probar que $f(\lambda)$ es un valor propio de $f(A)$ para todo polinomio $f(t)$.

8.51. Mostrar que dos matrices similares tienen los mismos valores propios.

8.52. Probar que las matrices A y A^T tienen los mismos valores propios. Dar un ejemplo en el que A y A^T tengan diferentes vectores propios.

POLINOMIOS CARACTERISTICO Y MINIMO

8.53. Hallar los polinomios característico y mínimo de cada una de las matrices:

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \qquad C = \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

8.54. Sean $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Mostrar que A y B tienen polinomios característicos diferentes (y por tanto no son similares) pero tienen el mismo polinomio mínimo. Así matrices no similares pueden tener el mismo polinomio mínimo.

8.55. Considérese una matriz cuadrada por bloques $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Probar que $tI - M = \begin{pmatrix} tI - A & -B \\ -C & tI - D \end{pmatrix}$ es la matriz característica de M.

8.56. Sea A una matriz n-cuadrada tal que $A^k = 0$ para algún $k > n$. Probar que $A^n = 0$.

8.57. Mostrar que una matriz A y su traspuesta A^T tienen el mismo polinomio mínimo.

8.58. Supóngase que $f(t)$ es un polinomio irreducible normalizado tal que $f(A) = 0$ para una matriz A. Mostrar que $f(t)$ es el polinomio mínimo de A.

8.59. Probar que A es una matriz escalar kI si y sólo si el polinomio mínimo de A es $m(t) = t - k$.

8.60. Hallar una matriz A cuyo polinomio mínimo sea *a*) $t^3 - 5t^2 + 6t + 8$, *b*) $t^4 - 5t^3 - 2t + 7t + 4$.

8.61. Considérense las matrices n-cuadradas escritas a continuación (donde $a \neq 0$):

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \qquad M = \begin{pmatrix} \lambda & a & 0 & \dots & 0 & 0 \\ 0 & \lambda & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & a \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Aquí N tiene 1 en la primera diagonal sobre la diagonal principal y 0 en cualquier otro sitio, y M tiene λ en la diagonal principal, a en la primera diagonal sobre la principal y 0 en el resto de las posiciones.

a) Probar que para $k < n$ N^k tiene 1 en la k-ésima diagonal sobre la principal y 0 en cualquier otra posición. Probar, asimismo, que $N^n = 0$.

b) Mostrar que el polinomio característico y mínimo de N es $f(t) = t^n$.

c) Mostrar que el polinomio característico y mínimo de M es $g(t) = (t - \lambda)^n$. (*Indicación:* Nótese que $M = \lambda I + aN$.)

DIAGONALIZACION

8.62. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz sobre el cuerpo real **R**. Determinar las condiciones necesarias y suficientes a imponer sobre a, b, c y d para que A sea diagonalizable, es decir, para que tenga dos vectores propios linealmente independientes.

8.63. Repetir el Problema 8.62 para el caso en el que A es una matriz sobre el cuerpo complejo **C**.

8.64. Probar que una matriz A es diagonalizable si y sólo si su polinomio mínimo es producto de factores lineales distintos entre sí.

8.65. Supóngase que E es una matriz tal que $E^2 = E$.

a) Encontrar el polinomio mínimo $m(t)$ de E.

b) Probar que E es diagonalizable y, más aún, E es similar a la matriz diagonal $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, donde r es el rango de E.

DIAGONALIZACION DE MATRICES REALES SIMETRICAS Y DE FORMAS CUADRATICAS

8.66. Para cada una de las matrices simétricas A siguientes, encontrar una matriz ortogonal P tal que $P^{-1}AP$ sea diagonal:

$$a)\ A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}, \qquad b)\ A = \begin{pmatrix} 5 & 4 \\ 4 & -1 \end{pmatrix}, \qquad c)\ A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$$

8.67. Hallar una transformación de coordenadas ortogonal que diagonalice cada forma cuadrática:

$$a)\ q(x, y) = 2x^2 - 6xy + 10y^2, \qquad b)\ q(x, y) = x^2 + 8xy - 5y^2$$

8.68. Hallar una transformación de coordenadas ortogonal que diagonalice la forma cuadrática $q(x, y, z) = 2xy + 2xz + 2yz$.

8.69. Sea A una matriz 2×2 real simétrica con valores propios 2 y 3. Sea $u = (1, 2)$ un vector propio perteneciente a 2. Encontrar un vector propio v perteneciente a 3 y A.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

8.40. $f(A) = \begin{pmatrix} -26 & -3 \\ 5 & -27 \end{pmatrix}$, $g(A) = \begin{pmatrix} -40 & 39 \\ -65 & -27 \end{pmatrix}$, $f(B) = \begin{pmatrix} 3 & 6 \\ 0 & 9 \end{pmatrix}$, $g(B) = \begin{pmatrix} 3 & 12 \\ 0 & 15 \end{pmatrix}$

8.41. $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$

8.42. *Indicación:* Sea $A = \begin{pmatrix} 2 & a & b \\ 0 & 2 & c \\ 0 & 0 & 2 \end{pmatrix}$. Hacer $B = A^3$ y después obtener condiciones sobre a, b y c.

8.45. *a*) $\lambda_1 = 1, u = (3, -2); \lambda_2 = 2, v = (2, -1)$ *b*) $P = \begin{pmatrix} -1 & -2 \\ 2 & 3 \end{pmatrix}$

c) $A^{10} = \begin{pmatrix} 4093 & 6138 \\ -2046 & -3066 \end{pmatrix}, f(A) = \begin{pmatrix} 2 & -6 \\ 2 & 9 \end{pmatrix}$ *d*) $B = \begin{pmatrix} -3 + 4\sqrt{2} & -6 + 6\sqrt{2} \\ 2 - 2\sqrt{2} & 4 - 3\sqrt{2} \end{pmatrix}$

8.46. *a)* $\lambda_1 = 2, u = (1, -1, 0), v = (1, 0, -1); \lambda_2 = 6, w = (1, 2, 1)$.

b) $\lambda_1 = 3, u = (1, 1, 0), v = (1, 0, 1); \lambda_2 = 1, w = (2, -1, 1)$.

c) $\lambda = 1, u = (1, 0, 0), v = (0, 0, 1)$.

Sean $P_1 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{pmatrix}$ y $P_2 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$. P_3 no existe porque C tiene a lo sumo dos vectores propios linealmente independientes y por tanto no puede diagonalizarse.

8.47. *a)* Para A, $\lambda = 3$, $u = (1, -1)$; B no tiene valores propios (en **R**).

b) Para A, $\lambda = 3$, $u = (1, -1)$; para B, $\lambda_1 = 2i$, $u = (1, 3 - 2i)$; $\lambda_2 = -2i$, $v = (1, 3 + 2i)$.

8.52. Sea $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Entonces $\lambda = 1$ es el único valor propio y $v = (1, 0)$ genera el espacio propio de $\lambda = 1$. Por otra parte, para $A^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\lambda = 1$ sigue siendo el único valor propio, pero $w = (0, 1)$ genera el espacio propio de $\lambda = 1$.

8.53. *a)* $\Delta(t) = (t - 2)^3(t - 7)^2$; $m(t) = (t - 2)^2(t - 7)$.

b) $\Delta(t) = (t - 3)^5$; $m(t) = (t - 3)^3$.

c) $\Delta(t) = (t - \lambda)^5$; $m(t) = t - \lambda$.

8.60. *a)* $A = \begin{pmatrix} 0 & 0 & -8 \\ 1 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}$, *b)* $A = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{pmatrix}$

8.65. *a)* Si $E = I$, $m(t) = (t - 1)$; si $E = 0$, $m(t) = t$; en cualquier otro caso, $m(t) = t(t - 1)$.

b) *Indicación:* Usar *a*).

8.66. *a)* $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, *b)* $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, *c)* $P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}$

8.67. *a)* $x = (3x' - y')/\sqrt{10}$, $y = (x' + 3y')/\sqrt{10}$, *b)* $x = (2x' - y')/\sqrt{5}$, $y = (x' + 2y')/\sqrt{5}$.

8.68. $x = x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6}$, $y = x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6}$, $z = x'/\sqrt{3} - 2z'/\sqrt{6}$.

8.69. $v = (2, -1)$, $A = \begin{pmatrix} \frac{14}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{11}{5} \end{pmatrix}$

CAPITULO 9

Aplicaciones lineales

9.1. INTRODUCCION

El principal objetivo del álgebra lineal lo constituyen los espacios vectoriales de dimensión finita y las aplicaciones lineales entre ellos. Los espacios vectoriales se presentaron en el Capítulo 5. Este capítulo introduce las aplicaciones lineales. No obstante, comenzaremos con una discusión de las aplicaciones en general.

9.2. APLICACIONES

Sea A y B conjuntos no vacíos arbitrarios. Supongamos que a cada elemento de A se le asigna un único elemento de B. La colección de tales asignaciones se denomina una *aplicación* de A en B. El conjunto A se llama el *dominio* de la aplicación y el conjunto B, su *codominio*. Una aplicación f de A en B se denota por

$$f: A \to B$$

Escribimos $f(a)$, leído «f de a», para representar el elemento de B que f asigna al elemento $a \in A$. Recibe el nombre de *valor* de f en a o *imagen* de a bajo f.

Nota: El término *función* se usará como sinónimo de aplicación, aunque en algunos textos se reserva la palabra función para designar las aplicaciones de valores reales o complejos, es decir, las que aplican un conjunto en **R** o **C**.

Consideremos una aplicación $f: A \to B$. Si A' es cualquier subconjunto de A, $f(A')$ denota el conjunto de imágenes de A'; y si B' es cualquier subconjunto de B, $f^{-1}(B')$ denota el conjunto de elementos de A cuyas imágenes están en B':

$$f(A') = \{f(a) : a \in A'\} \qquad \text{y} \qquad f^{-1}(B') = \{a \in A : f(a) \in B'\}$$

Llamamos a $f(A')$ la *imagen* de A' y a $f^{-1}(B')$ la *imagen inversa* o *preimagen* de B'. En particular, el conjunto de todas las imágenes, o sea, $f(A)$, se conoce como la imagen (o *recorrido*) de f.

A cada aplicación $f: A \to B$ le corresponde el subconjunto de $A \times B$ dado por $\{(a, f(a)): a \in A\}$. Este conjunto se denomina el *gráfico* de f. Se dice que dos aplicaciones $f: A \to B$ y $g: A \to B$ son *iguales*, escrito $f = g$, si $f(a) = g(a)$ para todo $a \in A$, esto es, si tienen el mismo gráfico. Así pues, no distinguiremos entre una función y su gráfico. La negación de $f = g$ se escribe $f \neq g$ y es la proposición:

$$\text{Existe un } a \in A \text{ para el cual } f(a) \neq g(a)$$

A veces la flecha «con barra» $\mapsto$ se utiliza para denotar la imagen de un elemento arbitrario $x \in A$ bajo una aplicación $f: A \to B$, escribiendo

$$x \mapsto f(x)$$

Esto se ilustra en el siguiente ejemplo.

EJEMPLO 9.1

a) Sea $f: \mathbf{R} \to \mathbf{R}$ la aplicación que asigna a cada número real x su cuadrado x^2:

$$x \mapsto x^2 \qquad \text{o} \qquad f(x) = x^2$$

La imagen de -3 es 9, de modo que escribimos $f(-3) = 9$.

b) Consideremos la matriz 2×3 $A = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix}$. Si expresamos los vectores en $\mathbf{R}^3$ y $\mathbf{R}^2$ como vectores columnas, A determina la aplicación $F: \mathbf{R}^3 \to \mathbf{R}^2$ definida por

$$v \mapsto Av \qquad \text{esto es} \qquad F(v) = Av \qquad v \in \mathbf{R}^3$$

Así, si $v = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$, entonces $F(v) = Av = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -10 \\ 12 \end{pmatrix}$.

c) Sea V el espacio vectorial de los polinomios en la variable t sobre el cuerpo real $\mathbf{R}$. La derivada define una aplicación $\mathbf{D}: V \to V$, donde para cada polinomio $f \in V$ hacemos $\mathbf{D}(f) = df/dt$. Por ejemplo, $\mathbf{D}(3t^2 - 5t + 2) = 6t - 5$.

d) Sea V el espacio vectorial de los polinomios en t sobre $\mathbf{R}$ [como en *c*)]. La integral, digamos 0 a 1, define una aplicación $\mathbf{J}: V \to \mathbf{R}$, donde para cada polinomio, $f \in V$, hacemos $\mathbf{J}(f) = \int_0^1 f(t)\, dt$. Por ejemplo,

$$\mathbf{J}(3t^2 - 5t + 2) = \int_0^1 (3t^2 - 5t + 2)\, dt = \tfrac{1}{2}$$

Nótese que esta aplicación es del espacio vectorial V en el cuerpo real $\mathbf{R}$, mientras que la de *c*) es de V en sí mismo.

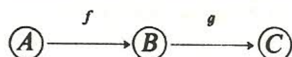
Nota: Toda matriz $m \times n$ sobre un cuerpo K determina la aplicación $F: K^n \to K^m$ definida por

$$v \mapsto Av$$

donde los vectores en K^n y K^m se escriben como vectores columna. Por razones de conveniencia, denotaremos usualmente la aplicación precedente por A, el mismo símbolo empleado para la matriz.

COMPOSICION DE APLICACIONES

Consideremos dos aplicaciones $f: A \to B$ y $g: B \to C$ como se ilustra abajo:



Sea $a \in A$. Entonces $f(a) \in B$, que es el dominio de g. Por tanto, podemos obtener la imagen de $f(a)$ bajo la aplicación g, es decir, $g(f(a))$. Esta aplicación

$$a \mapsto g(f(a))$$

de A en C se llama la *composición* de f y g y se denota por $g \circ f$. En otras palabras, $(g \circ f): A \to C$ es la aplicación definida por

$$(g \circ f)(a) = g(f(a))$$

Nuestro primer teorema nos dice que la composición de aplicaciones satisface la ley asociativa.

Teorema 9.1: Sean $f: A \to B$, $g: B \to C$ y $h: C \to D$. En tal caso, $h \circ (g \circ f) = (h \circ g) \circ f$.

Demostremos ahora este teorema. Si $a \in A$,

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a)))$$

y

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a)))$$

De este modo, $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$ para todo $a \in A$, luego $h \circ (g \circ f) = (h \circ g) \circ f$.

Nota: Sea $F: A \to B$. En algunos textos se escribe aF en lugar de $F(a)$ para designar la imagen de $a \in A$ bajo F. Con esta notación, la composición de las funciones $F: A \to B$ y $G: B \to C$ se expresa $F \circ G$ y no $G \circ F$ como se hace en este texto.

APLICACIONES INYECTIVAS (UNO-A-UNO) Y SUPRAYECTIVAS (SOBRE)

Introducimos formalmente algunos tipos especiales de aplicaciones.

Definición: Se dice que una aplicación $f: A \to B$ es *inyectiva* o *uno-a-uno* (también *uno-uno* o 1-1) si los elementos diferentes de A tienen imágenes diferentes; esto es, si

$$a \neq a' \quad \text{implica} \quad f(a) \neq f(a')$$

o, equivalentemente,

$$f(a) = f(a') \quad \text{implica} \quad a = a'$$

Definición: Se dice que una aplicación $f: A \to B$ es *suprayectiva* o *sobre* (o que f aplica A *sobre* B) si todo $b \in B$ es la imagen de, al menos, un $a \in A$.

Se dice que una aplicación $f: A \to B$ es *biyectiva*, o que es una *correspondencia uno-a-uno entre* A y B, si es simultáneamente inyectiva y suprayectiva.

EJEMPLO 9.2

a) Definamos $f: \mathbf{R} \to \mathbf{R}$, $g: \mathbf{R} \to \mathbf{R}$ y $h: \mathbf{R} \to \mathbf{R}$ según $f(x) = 2^x$, $g(x) = x^3 - x$ y $h(x) = x^2$. Los gráficos de estas funciones se muestran en la Figura 9-1. La aplicación f es inyectiva. Geométricamente, esto significa que cada recta horizontal no contiene más de un punto de f. La aplicación g es suprayectiva. Geométricamente, esto quiere decir que cada recta horizontal contiene al menos un punto de g. La aplicación h no es ni inyectiva ni suprayectiva; por ejemplo, 2 y -2 tienen la misma imagen, 4 y -16 no es imagen de ningún elemento de $\mathbf{R}$.

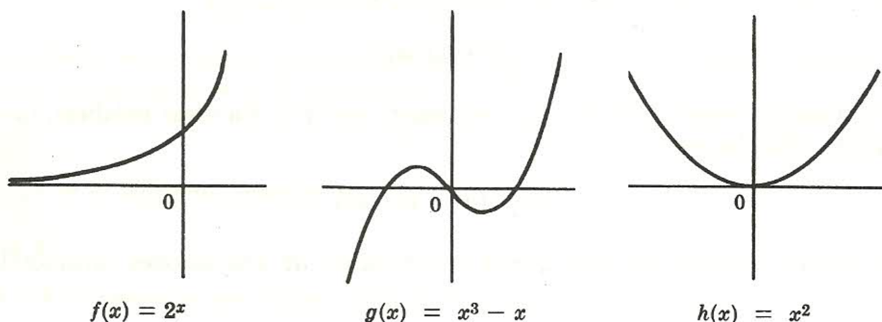


Figura 9-1.

b) Sea A cualquier conjunto. La aplicación $f: A \to A$ definida por $f(a) = a$, es decir, la que asigna a cada elemento de A el propio elemento, se denomina *aplicación identidad* en A y se denota por 1_A, 1 o I.

c) Sea $f: A \to B$. Llamamos a $g: B \to A$ la inversa de f, escrito f^{-1}, si

$$f \circ g = 1_B \quad \text{y} \quad g \circ f = 1_A$$

Señalamos que f tiene una inversa si y sólo si es tanto inyectiva como suprayectiva (Problema 9.11). Además, si $b \in B$, $f^{-1}(b) = a$, siendo a el único elemento de A para el que $f(a) = b$.

9.3. APLICACIONES LINEALES

Sea U y V espacios vectoriales sobre el mismo cuerpo K. Una aplicación $F: V \to U$ se llama *aplicación lineal* (o *transformación lineal* u *homomorfismo entre espacios vectoriales*) si satisface las dos condiciones:

1. Para todo par $v, w \in V$, $F(v + w) = F(v) + F(w)$.
2. Para todo $k \in K$ y todo $v \in V$, $F(kv) = kF(v)$.

Dicho de otro modo, $F: V \to U$ es lineal si «preserva» las dos operaciones básicas de un espacio vectorial, la de suma vectorial y la de producto por un escalar.

Sustituyendo $k = 0$ en 2 obtenemos $F(0) = 0$. Esto es, toda aplicación lineal lleva el vector cero al vector cero.

Ahora, para todo par de escalares $a, b \in K$ y todo par de vectores $v, w \in V$ obtenemos, imponiendo ambas condiciones de linealidad,

$$F(av + bw) = F(av) + F(bw) = aF(v) + bF(w)$$

Con mayor generalidad, para escalares cualesquiera $a_i \in K$ y vectores cualesquiera $v_i \in V$ llegamos a la propiedad básica de las aplicaciones lineales:

$$F(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 F(v_1) + a_2 F(v_2) + \cdots + a_n F(v_n)$$

Nota: La condición $F(av + bw) = aF(v) + bF(w)$ caracteriza completamente las aplicaciones lineales y se utiliza a veces para definirlas.

EJEMPLO 9.3

a) Sea A cualquier matriz $m \times n$ sobre un cuerpo K. Como se señaló previamente, A determina una aplicación $F: K^n \to K^m$ mediante la asignación $v \mapsto Av$. (Aquí los vectores de K^n y K^m se escriben como columnas.) Afirmamos que F es lineal. Esto es debido a que, por propiedades de las matrices,

$$F(v + w) = A(v + w) = Av + Aw = F(v) + F(w)$$

y

$$F(kv) = A(kv) = kAv = kF(v)$$

donde $v, w \in K^n$ y $k \in K$.

b) Sea $F: \mathbf{R}^3 \to \mathbf{R}^3$ la aplicación «proyección» en el plano xy: $F(x, y, z) = (x, y, 0)$. Probemos que F es lineal. Sean $v = (a, b, c)$ y $w = (a', b', c')$. Entonces

$$F(v + w) = F(a + a', b + b', c + c') = (a + a', b + b', 0) =$$
$$= (a, b, 0) + (a', b', 0) = F(v) + F(w)$$

y, para todo $k \in \mathbf{R}$,

$$F(kv) = F(ka, kb, kc) = (ka, kb, 0) = k(a, b, 0) = kF(v)$$

O sea, F es lineal.

c) Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación de «traslación» definida según $F(x, y) = (x + 1, y + 2)$. Obsérvese que $F(0) = F(0, 0) = (1, 2) \neq 0$. Es decir, el vector cero no se aplica sobre el vector cero. Por consiguiente, F no es lineal.

d) Sea $F: V \to U$ la aplicación que asigna $0 \in U$ a todo $v \in V$. Para todo par de vectores $v, w \in V$ y todo $k \in K$ tenemos

$$F(v + w) = 0 = 0 + 0 = F(v) + F(w) \qquad \text{y} \qquad F(kv) = 0 = k0 = kF(v)$$

Así F es lineal. Llamamos a F la *aplicación cero* y la denotaremos normalmente por **0**.

e) Consideremos la aplicación identidad $I: V \to V$ que aplica cada $v \in V$ en sí mismo. Para todo par de vectores $v, w \in V$ y todo par de escalares $a, b \in K$,

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Así I es lineal.

f) Sea V el espacio vectorial de los polinomios en la variable t sobre el cuerpo real $\mathbf{R}$. La aplicación derivada $\mathbf{D}: V \to V$ y la aplicación integral $\mathbf{J}: V \to \mathbf{R}$, definidas en el Ejemplo 9.1 *c*) y *d*), son lineales. La razón es que, según se demuestra en el cálculo, para todo par de vectores $u, v \in V$ y todo $k \in \mathbf{R}$,

$$\frac{d(u + v)}{dt} = \frac{du}{dt} + \frac{dv}{dt} \qquad \text{y} \qquad \frac{d(ku)}{dt} = k\,\frac{du}{dt}$$

esto es, $\mathbf{D}(u + v) = \mathbf{D}(u) + \mathbf{D}(v)$ y $\mathbf{D}(ku) = k\,\mathbf{D}(u)$; y también,

$$\int_0^1 [u(t) + v(t)]\,dt = \int_0^1 u(t)\,dt + \int_0^1 v(t)\,dt$$

e
$$\int_0^1 ku(t)\,dt = k \int_0^1 u(t)\,dt$$

esto es, $\mathbf{J}(u + v) = \mathbf{J}(u) + \mathbf{J}(v)$ y $\mathbf{J}(ku) = k\mathbf{J}(u)$.

g) Sea $F: V \to U$ una aplicación lineal que es simultáneamente inyectiva y suprayectiva. En tal caso, existe una aplicación inversa $F^{-1}: U \to V$. Probaremos (Problema 9.15) que esta aplicación inversa es también lineal.

Nuestro próximo teorema (demostrado en el Problema 9.12) nos proporciona abundantes ejemplos de aplicaciones lineales. En particular, nos dice que una aplicación lineal queda completamente determinada por los valores que asigna a los elementos de una base.

Teorema 9.2: Sean V y U espacios vectoriales sobre un cuerpo K. Sean $(v_1, v_2, \ldots, v_n)$ una base de V y $u_1, u_2, \ldots, u_n$ vectores cualesquiera en U. Existe una única aplicación lineal $F: V \to U$ tal que $F(v_1) = u_1$, $F(v_2) = u_2$, $\ldots$, $F(v_n) = u_n$.

Subrayamos el hecho de que los vectores $u_1, \ldots, u_n$ del Teorema 9.2 son completamente arbitrarios; pueden ser linealmente dependientes o incluso iguales entre sí.

ISOMORFISMOS ENTRE ESPACIOS VECTORIALES

La noción de espacios vectoriales isomorfos se definió en el Capítulo 5, cuando estudiábamos las coordenadas de un vector relativas a una base. Redefinimos ahora este concepto.

Definición: Dos espacios vectoriales V y U sobre K son *isomorfos* si existe una aplicación lineal biyectiva $F: V \to U$. La aplicación F se denomina entonces *isomorfismo* entre V y U.

EJEMPLO 9.4. Sean V un espacio sobre K de dimensión n y S una base de V. Como se indicó previamente, la aplicación $v \mapsto [v]_S$, que aplica cada $v \in V$ en su vector coordenado respecto a la base S, es un isomorfismo entre V y K^n.

9.4. NUCLEO E IMAGEN DE UNA APLICACION LINEAL

Empezamos definiendo dos conceptos.

Definición: Sea $F: V \to U$ una aplicación lineal. La *imagen* de F, escrito Im F, es el conjunto de puntos imagen en U:

$$\text{Im } F = \{u \in U : F(v) = u \text{ para algún } v \in V\}$$

El *núcleo* de F, escrito Ker F, es el conjunto de elementos de V que se aplican en $0 \in U$:

$$\text{Ker } F = \{v \in V : F(v) = 0\}$$

Se demuestra fácilmente (Problema 9.22) el teorema enunciado a continuación.

Teorema 9.3: Sea $F: V \to U$ una aplicación lineal. La imagen de F es un subespacio de U y el núcleo de F es un subespacio de V.

Supongamos ahora que los vectores $v_1, \ldots, v_n$ generan V y que $F: V \to U$ es lineal. Probemos que los vectores $F(v_1), \ldots, F(v_n) \in U$ generan Im F. Supongamos que $u \in \text{Im } F$; en ese caso, $F(v) = u$ para algún vector $v \in V$. Como los v_i generan V y $v \in V$, existen escalares $a_1, \ldots, a_n$ tales que

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

En consecuencia,

$$u = F(v) = F(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 F(v_1) + a_2 F(v_2) + \cdots + a_n F(v_n)$$

y por tanto los vectores $F(v_1), \ldots, F(v_n)$ generan Im F.

Establezcamos formalmente este útil resultado.

Proposición 9.4: Supongamos que $v_1, v_2, \ldots, v_n$ generan un espacio vectorial V y que $F: V \to U$ es lineal. Entonces $F(v_1), F(v_2), \ldots, F(v_n)$ generan Im F.

EJEMPLO 9.5

a) Sea $F: \mathbf{R}^3 \to \mathbf{R}^3$ la aplicación proyección sobre el plano xy. Esto es,

$$F(x, y, z) = (x, y, 0)$$

(Véase la Figura 9-2.) Claramente la imagen de F es todo el plano xy. Es decir,

$$\text{Im } F = \{(a, b, 0) : a, b \in \mathbf{R}\}$$

Nótese que el núcleo de F es el eje z. O sea,

$$\text{Ker } F = \{(0, 0, c) : c \in \mathbf{R}\}$$

ya que estos puntos y solamente éstos se aplican en el vector cero $0 = (0, 0, 0)$.

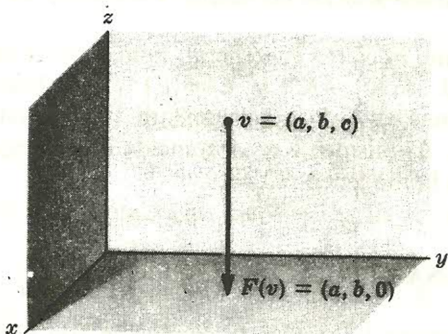


Figura 9-2.

b) Sean V el espacio vectorial de los polinomios sobre $\mathbf{R}$ y $\mathbf{T}: V \to V$ el operador tercera derivada, esto es,

$$\mathbf{T}[f(t)] = d^3f/dt^3$$

[A veces se emplea la notación $\mathbf{T} = \mathbf{D}^3$, donde $\mathbf{D}$ es la aplicación derivada del Ejemplo 9.1 *c*).] Tendremos

$$\text{Ker } \mathbf{T} = \{\text{polinomios de grado} \leq 2\}$$

[porque $\mathbf{T}(at^2 + bt + c) = 0$ pero $\mathbf{T}(t^n) \neq 0$ para $n > 3$]. Por otra parte,

$$\text{Im } \mathbf{T} = V$$

puesto que todo polinomio en V es la tercera derivada de algún polinomio.

c) Consideremos una matriz 4×3 arbitraria A sobre un cuerpo K:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}$$

que vemos como aplicación lineal $A: K^3 \to K^4$. La base usual $\{e_1, e_2, e_3\}$ de K^3 genera K^3, luego sus valores Ae_1, Ae_2, Ae_3 bajo A generan la imagen de A. Pero los vectores Ae_1, Ae_2 y Ae_3 son las columnas de A:

$$Ae_1 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix}, \qquad Ae_2 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix}$$

$$Ae_3 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \\ d_3 \end{pmatrix}$$

de modo que la imagen de A es precisamente el espacio columna de A.

Por otra parte, el núcleo de A consiste en todos los vectores v para los que $Av = 0$. Esto significa que el núcleo de A es el espacio solución del sistema homogéneo $AX = 0$.

Nota: El resultado precedente es cierto en general. Esto es, si A es cualquier matriz $m \times n$ vista como una aplicación lineal $A: K^n \to K^m$ y $E = \{e_i\}$ es la base usual de K^n, los vectores $Ae_1, \ldots, Ae_n$ serán las columnas de A. Por consiguiente, Im A = c-lin A, que quiere decir espacio columna de A (Sección 5.5). Asimismo, Ker A coincide con el *espacio nulo* de A, o sea, con el espacio solución del sistema homogéneo $AX = 0$.

RANGO Y NULIDAD DE UNA APLICACION LINEAL

Hasta aquí no hemos relacionado la noción de dimensión con la de aplicación lineal $F: V \to U$. En los casos en que V es de dimensión finita, tenemos la relación fundamental que sigue.

Teorema 9.5: Sea V de dimensión finita y sea $F: V \to U$ una aplicación lineal. Entonces

$$\dim V = \dim(\operatorname{Ker} F) + \dim(\operatorname{Im} F) \qquad [9.1]$$

Es decir, la suma de las dimensiones de la imagen y el núcleo de una aplicación lineal es igual a la dimensión de su dominio.

Se ve fácilmente que la ecuación [9.1] rige para la aplicación proyección F del Ejemplo 9.5 *a*). Allí la imagen (plano xy) y el núcleo (eje z) de F tienen dimensiones 2 y 1, respectivamente, mientras que el dominio $\mathbf{R}^3$ de F tiene dimensión 3.

Nota: Sea $F: V \to U$ una aplicación lineal. Se define el *rango* de F como la dimensión de su imagen y la *nulidad* de F como la dimensión de su núcleo; esto es,

$$\text{rango } F = \dim(\operatorname{Im} F) \qquad \text{y} \qquad \text{nulidad } F = \dim(\operatorname{Ker} F)$$

El Teorema 9.5 proporciona, pues, la siguiente fórmula para F cuando V tiene dimensión finita:

$$\text{rango } F + \text{nulidad } F = \dim V$$

Recordemos que el rango de una matriz se definió en origen como la dimensión de su espacio columna y de su espacio fila. Obsérvese que si vemos ahora A como una aplicación lineal, las dos definiciones se corresponden, porque la imagen de A es precisamente su espacio columna.

EJEMPLO 9.6. Sea $F: \mathbf{R}^4 \to \mathbf{R}^3$ la aplicación lineal definida por

$$F(x, y, s, t) = (x - y + s + t, x + 2s - t, x + y + 3s - 3t)$$

a) Encontremos una base y la dimensión de la imagen de F.

Hallamos la imagen de los vectores de la base usual de $\mathbf{R}^4$:

$$F(1, 0, 0, 0) = (1, 1, 1) \qquad F(0, 0, 1, 0) = (1, 2, 3)$$

$$F(0, 1, 0, 0) = (-1, 0, 1) \qquad F(0, 0, 0, 1) = (1, -1, -3)$$

Según la Proposición 9.4, los vectores imagen generan Im F; por eso construimos la matriz cuyas filas son estos vectores imagen y la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

De este modo, $(1, 1, 1)$ y $(0, 1, 2)$ constituyen una base de Im F, luego $\dim(\operatorname{Im} F) = 2$ o, en otras palabras, rango $F = 2$.

b) Encontremos una base y la dimensión del núcleo de la aplicación F.

Hacemos $F(v) = 0$, donde $v = (x, y, z, t)$:

$$F(x, y, s, t) = (x - y + s + t, x + 2s - t, x + y + 3s - 3t) = (0, 0, 0)$$

Igualamos entre sí las componentes correspondientes para formar el siguiente sistema homogéneo cuyo espacio solución es Ker F:

$$\begin{aligned} x - y + s + t &= 0 \\ x \quad + 2s - t &= 0 \\ x + y + 3s - 3t &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \\ 2y + 2s - 4t &= 0 \end{aligned} \quad \text{o} \quad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \end{aligned}$$

Las variables libres son s y t, luego dim (Ker F) = 2 o nulidad $F = 2$. Tomamos:

i) $s = -1$, $t = 0$, para obtener la solución $(2, 1, -1, 0)$.

ii) $s = 0$, $t = 1$, para obtener la solución $(1, 2, 0, 1)$.

Así $(2, 1, -1, 0)$ y $(1, 2, 0, 1)$ constituyen una base de Ker F. (Obsérvese que rango F + nulidad $F = 2 + 2 = 4$, que es la dimensión del dominio $\mathbf{R}^4$ de F.)

APLICACION A LOS SISTEMAS DE ECUACIONES LINEALES

Consideremos un sistema de m ecuaciones lineales con n incógnitas sobre un cuerpo K:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

que es equivalente a la ecuación matricial

$$Ax = b$$

donde $A = (a_{ij})$ es la matriz de los coeficientes y $x = (x_i)$ y $b = (b_i)$ los vectores columna de las incógnitas y de las constantes, respectivamente. Ahora la matriz A puede verse también como la aplicación lineal

$$A : K^n \to K^m$$

Siendo así, la solución de la ecuación $Ax = b$ puede verse como la preimagen de $b \in K^m$ bajo la aplicación lineal $A : K^n \to K^m$. Más aún, la solución de la ecuación homogénea asociada $Ax = 0$ puede verse como el núcleo de la aplicación lineal $A : K^n \to K^m$.

El Teorema 9.5 relativo a aplicaciones lineales nos conduce a la relación:

$$\dim(\text{Ker } A) = \dim K^n - \dim(\text{Im } A) = n - \text{rango } A$$

Pero n es exactamente el número de incógnitas en el sistema homogéneo $Ax = 0$, por lo que volvemos a encontrar el Teorema 5.20.

Teorema 9.6: La dimensión del espacio solución W del sistema de ecuaciones lineales homogéneo $AX = 0$ es $n - r$, donde n es el número de incógnitas y r el rango de la matriz de los coeficientes A.

9.5. APLICACIONES LINEALES SINGULARES Y NO SINGULARES. ISOMORFISMOS

Se dice que una aplicación lineal $F: V \to U$ es *singular* si la imagen de algún vector no nulo bajo F es 0, es decir, si existe algún $v \in V$ tal que $v \neq 0$ pero $F(v) = 0$. De esta menera, $F: V \to U$ es *no singular* si únicamente $0 \in V$ se aplica en $0 \in U$ o, equivalentemente, si su núcleo consiste solamente en el vector cero: $\text{Ker } F = \{0\}$.

Una propiedad fundamental de las aplicaciones no singulares es la que ahora escribimos (véase la demostración en el Problema 9.29).

Teorema 9.7: Supongamos que una aplicación lineal $F: V \to U$ es no singular. En tal caso, la imagen de cualquier conjunto linealmente independiente es linealmente independiente.

ISOMORFISMOS

Supongamos que una aplicación lineal $F: V \to U$ es inyectiva. Entonces sólo $0 \in V$ puede aplicarse en $0 \in U$ y por tanto F es no singular. El recíproco es cierto también, porque si F es no singular y $F(v) = F(w)$, necesariamente $F(v - w) = F(v) - F(w) = 0$, luego $v - w = 0$ o $v = w$. Así $F(v) = F(w)$ implica $v = w$, o sea, F es inyectiva. Hemos demostrado, pues,

Proposición 9.8: Una aplicación lineal $F: V \to U$ es inyectiva si y sólo si es no singular.

Recordemos que una aplicación $F: V \to U$ recibe el nombre de isomorfismo si es lineal y biyectiva, esto es, si es lineal, inyectiva y suprayectiva. Recordemos, asimismo, que se dice que un espacio vectorial V es isomorfo a otro U, escrito $V \simeq U$, si existe algún isomorfismo $F: V \to U$.

Es aplicable el teorema enunciado a continuación, demostrado en el Problema 9.30.

Teorema 9.9: Supongamos que V tiene dimensión finita y que $\dim V = \dim U$. Supongamos también que $F: V \to U$ es lineal. Entonces F es un isomorfismo si y sólo si es no singular.

9.6. OPERACIONES CON APLICACIONES LINEALES

Podemos combinar aplicaciones lineales de varias maneras consiguiendo nuevas aplicaciones lineales. Estas operaciones son de gran importancia y se utilizarán a lo largo de todo el texto.

Supongamos que $F: V \to U$ y $G: V \to U$ son aplicaciones lineales entre espacios vectoriales sobre un cuerpo K. Definimos la suma $F + G$ como la aplicación de V en U que asigna $F(v) + G(v)$ a $v \in V$:

$$(F + G)(v) = F(v) + G(v)$$

Asimismo, para cualquier escalar $k \in K$, definimos el producto kF como la aplicación de V en U que asigna $kF(v)$ a $v \in V$:

$$(kF)(v) = kF(v)$$

Probamos ahora que si F y G son lineales, también lo son $F + G$ y kF. Tenemos, para todo par de vectores v, $w \in V$ y todo par de escalares a, $b \in K$,

$$\begin{aligned}(F + G)(av + bw) &= F(av + bw) + G(av + bw) = \\ &= aF(v) + bF(w) + aG(v) + bG(w) = \\ &= a(F(v) + G(v)) + b(F(w) + G(w)) = \\ &= a(F + G)(v) + b(F + G)(w) =\end{aligned}$$

y

$$\begin{aligned}(kF)(av + bw) &= kF(av + bw) = k(aF(v) + bF(w)) = \\ &= akF(v) + bkF(w) = a(kF)(v) + b(kF)(w)\end{aligned}$$

Así pues, $F + G$ y kF son lineales.

Disponemos del siguiente teorema.

Teorema 9.10: Sean V y U espacios vectoriales sobre un cuerpo K. La colección de todas las aplicaciones lineales de V en U, con las operaciones de suma y producto por un escalar anteriores, es un espacio vectorial sobre K.

El espacio vectorial del Teorema 9.10 suele denotarse por

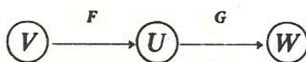
$$\text{Hom}\ (V, U)$$

Aquí Hom viene de la palabra homomorfismo. En caso de ser V y U de dimensión finita, se aplica el teorema que enseguida enunciamos, demostrado en el Problema 9.36.

Teorema 9.11: Supongamos $\dim V = m$ y $\dim U = n$. En tal caso, $\dim \text{Hom}(V, U) = mn$.

COMPOSICION DE APLICACIONES LINEALES

Sean V, U y W espacios vectoriales sobre un mismo cuerpo K y $F: V \to U$ y $G: U \to W$ aplicaciones lineales:



Recordemos que la función compuesta $G \circ F$ es la aplicación de V en W definida por $(G \circ F)(v) = G(F(v))$. Probemos que $G \circ F$ es lineal siempre que lo sean F y G. Tenemos, para todo par de vectores v, $w \in V$ y todo par de escalares a, $b \in K$,

$$\begin{aligned}(G \circ F)(av + bw) &= G(F(av + bw)) = G(aF(v) + bF(w)) = \\ &= aG(F(v)) + bG(F(w)) = a(G \circ F)(v) + b(G \circ F)(w)\end{aligned}$$

Es decir, $G \circ F$ es lineal.

La composición de aplicaciones lineales está relacionada con la suma y el producto por un escalar como sigue (véase la demostración en el Problema 9.37):

Teorema 9.12: Sean V, U y W espacios vectoriales sobre K, F y F' aplicaciones lineales de V en U, G y G' aplicaciones lineales de U en W y sea $k \in K$. Entonces

i) $G \circ (F + F') = G \circ F + G \circ F'$.

ii) $(G + G') \circ F = G \circ F + G' \circ F$.

iii) $k(G \circ F) = (kG) \circ F = G \circ (kF)$.

9.7. ALGEBRA DE OPERADORES LINEALES $A(V)$

Sea V un espacio vectorial sobre un cuerpo K. Consideramos ahora el caso especial de las aplicaciones lineales $F: V \to V$, o sea, de V en sí mismo. Estas se llaman también *operadores lineales* o *transformaciones lineales* en V. Escribiremos $A(V)$, en lugar de $\text{Hom}(V, V)$, para designar el espacio de tales aplicaciones.

De acuerdo con el Teorema 9.10, $A(V)$ es un espacio vectorial sobre K; es de dimensión n^2 si V es de dimensión n. Si $F, G \in A(V)$, la composición $G \circ F$ existe y es también una aplicación lineal de V en sí mismo, es decir, $G \circ F \in A(V)$. De este modo, tenemos un «producto» definido en $A(V)$. [En el espacio $A(V)$ expresaremos $G \circ F$ como GF.]

Nota: Un álgebra A sobre un cuerpo K es un espacio vectorial sobre K en el que se ha definido una operación de producto que satisface, para todos los $F, G, H \in A$ y todo $k \in K$,

i) $F(G + H) = FG + FH$

ii) $(G + H)F = GF + HF$

iii) $k(GF) = (kG)F = G(kF)$

Si la ley asociativa rige para el producto, esto es, si para todos los $F, G, H \in A$,

iv) $(FG)H = F(GH)$

se dice que el álgebra A es *asociativa*.

La definición precedente de álgebra y los Teoremas 9.10, 9.11 y 9.12 nos proporcionan el siguiente resultado básico.

Teorema 9.13: Sea V un espacio vectorial sobre K. En ese caso, $A(V)$ es un álgebra asociativa sobre K con respecto a la composición de aplicaciones. Si $\dim V = n$, $\dim A(V) = n^2$.

En vista del teorema anterior, $A(V)$ se denomina frecuentemente el *álgebra de operadores lineales* en V.

POLINOMIOS Y OPERADORES LINEALES

Observemos que la aplicación identidad $I: V \to V$ pertenece a $A(V)$. Además, para todo $T \in A(V)$, tenemos $TI = IT = T$. Hacemos notar que también pueden formarse «potencias» de T; usamos la notación $T^2 = T \circ T$, $T^3 = T \circ T \circ T, \ldots$. Más aún, para todo polinomio

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \qquad a_i \in K$$

podemos construir el operador $p(T)$ definido según

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n$$

(Para un escalar $k \in K$, el operador kI se denota con frecuencia por k, simplemente.) En particular, si $p(T) = \mathbf{0}$, la aplicación cero, se dice que T es un *cero* del polinomio $p(x)$.

EJEMPLO 9.7. Definamos $T: \mathbf{R}^3 \to \mathbf{R}^3$ por $T(x, y, z) = (0, x, y)$. Si (a, b, c) es cualquier elemento de $\mathbf{R}^3$,

$$(T + I)(a, b, c) = (0, a, b) + (a, b, c) = (a, a + b, b + c)$$

y

$$T^3(a, b, c) = T^2(0, a, b) = T(0, 0, a) = (0, 0, 0)$$

Vemos, pues, que $T^3 = \mathbf{0}$, la aplicación cero de V en sí mismo. Dicho de otro modo, T es un cero del polinomio $p(x) = x^3$.

9.8. OPERADORES INVERTIBLES

Decimos que un operador lineal $T: V \to V$ es *invertible* si tiene una inversa, es decir, si existe $T^{-1} \in A(V)$ tal que $TT^{-1} = T^{-1}T = I$.

Ahora bien, T es invertible si y sólo si es inyectivo y suprayectivo. De este modo, en particular, si T es invertible, solamente $0 \in V$ puede aplicarse en sí mismo, esto es, T es no singular. El recíproco no es cierto en general como se ve en el siguiente ejemplo.

EJEMPLO 9.8. Sean V el espacio vectorial de los polinomios sobre K y T el operador en V definido según

$$T(a_0 + a_1 t + \cdots + a_n t^n) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1} \qquad (n = 0, 1, 2, \ldots)$$

o sea, T aumenta en una unidad el exponente de t en cada término. Tenemos que T es una aplicación lineal y es no singular. Sin embargo, T no es suprayectivo y por tanto no es invertible.

La situación cambia significativamente cuando V es de dimensión finita. En concreto disponemos del teorema enunciado a continuación.

Teorema 9.14: Supongamos que T es un operador lineal en un espacio vectorial V de dimensión finita. Son equivalentes las cuatro condiciones:

i) T es no singular, es decir, $\operatorname{Ker} T = \{0\}$.
ii) T es inyectivo.
iii) T es suprayectivo.
(iv) T es invertible, esto es, inyectivo y suprayectivo.

La Proposición 9.8 nos dice que i) y ii) son equivalentes. Siendo así, para demostrar el teorema, sólo necesitamos probar que lo son i) y iii). [Se seguirá entonces que iv) es equivalente a las otras.] Según el Teorema 9.6,

$$\dim V = \dim (\operatorname{Im} T) + \dim (\operatorname{Ker} T)$$

Si T es no singular, necesariamente $\dim (\operatorname{Ker} T) = 0$, con lo cual, $\dim V = \dim (\operatorname{Im} T)$. Esto significa que $V = \operatorname{Im} T$ o que T es suprayectivo, de manera que i) implica iii). Recíprocamente,

supongamos que T es suprayectivo. En ese caso, $V = \text{Im } T$, de donde $\dim V = \dim (\text{Im } T)$. Esto quiere decir que $\dim (\text{Ker } T) = 0$ y por ende T es no singular. Así pues, iii) implica i). En consecuencia, el teorema queda demostrado. (La demostración del Teorema 9.9 es idéntica.)

EJEMPLO 9.9. Sea T el operador en $\mathbf{R}^2$ definido por $T(x, y) = (y, 2x - y)$. El núcleo de T es $\{(0, 0)\}$, luego T es no singular y, en virtud del Teorema 9.14, invertible. Hallemos ahora una fórmula para T^{-1}. Supongamos que (s, t) es la imagen de (x, y) bajo T; por consiguiente, (x, y) es la imagen de (s, t) bajo T^{-1}: $T(x, y) = (s, t)$ y $T^{-1}(s, t) = (x, y)$. Tenemos

$$T(x, y) = (y, 2x - y) = (s, t) \qquad \text{y así} \qquad y = s,\ 2x - y = t$$

Despejando x e y en términos de s y t obtenemos $x = \frac{1}{2}s + \frac{1}{2}t$, $y = s$. Entonces T^{-1} viene dado por la fórmula $T^{-1}(s, t) = (\frac{1}{2}s + \frac{1}{2}t, s)$.

APLICACIONES A LOS SISTEMAS DE ECUACIONES LINEALES

Consideremos un sistema de ecuaciones lineales sobre K y supongamos que el sistema tiene el mismo número de ecuaciones que de incógnitas, digamos n. Podemos representar este sistema por la ecuación matricial

$$Ax = b \tag{$*$}$$

donde A es una matriz n-cuadrada sobre K que vemos como un operador lineal en K^n. Supongamos que la matriz A es no singular, esto es, que la ecuación matricial $Ax = 0$ tiene sólo la solución trivial. La aplicación lineal A es, pues, inyectiva y suprayectiva. Esto significa que el sistema $(*)$ tiene solución única para cualquier $b \in K^n$. Por otra parte, supongamos que la matriz A es singular, es decir, que la ecuación matricial $Ax = 0$ tiene una solución no nula. En ese caso, la aplicación lineal A no es suprayectiva, lo que significa que existe $b \in K^n$ para el que $(*)$ no tiene solución. Más aún, si existe una solución, ésta no es única. [Véase el Problema 1.63 ii).] Hemos demostrado así el resultado fundamental:

Teorema 9.15: Consideremos el siguiente sistema de ecuaciones lineales con el mismo número de ecuaciones que de incógnitas:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}$$

a) Si el sistema homogéneo asociado sólo tiene la solución nula, el sistema precedente tiene solución única para todos los valores de los b_i.

b) Si el sistema homogéneo asociado tiene una solución no nula, entonces: i) hay valores de los b_i para los que el sistema precedente no tiene solución; ii) siempre que exista una solución del sistema anterior, no será única.

PROBLEMAS RESUELTOS

APLICACIONES

9.1. Decir si cada diagrama de la Figura 9-3 define o no una aplicación de $A = \{a, b, c\}$ en $B = \{x, y, z\}$.

a) No. Al elemento $b \in A$ no se le asigna nada.

b) No. A $c \in A$ se le asignan dos elementos, x y z.

c) Sí.

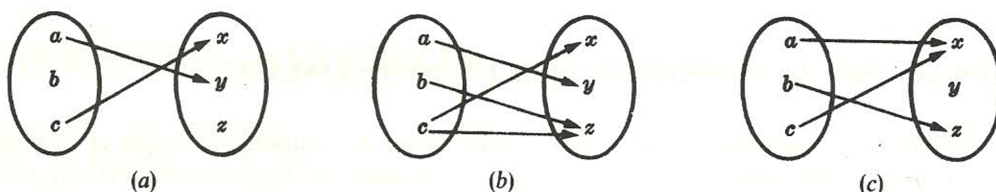


Figura 9-3.

9.2. Defínanse las aplicaciones $f: A \to B$ y $g: B \to C$ según el diagrama de la Figura 9-4. Hallar: *a*) la aplicación compuesta $(g \circ f): A \to C$, *b*) las imágenes de las aplicaciones f, g y $g \circ f$.

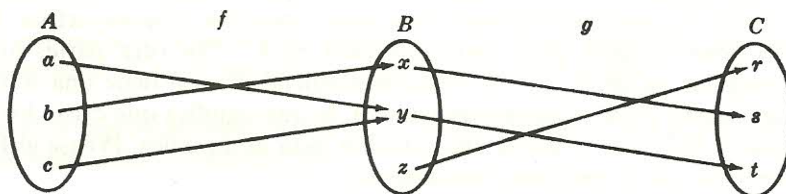


Figura 9-4.

a) Utilizamos la definición de aplicación compuesta para calcular:

$$(g \circ f)(a) = g(f(a)) = g(y) = t$$
$$(g \circ f)(b) = g(f(b)) = g(x) = s$$
$$(g \circ f)(c) = g(f(c)) = g(y) = t$$

Obsérvese que llegamos al mismo resultado si «seguimos las flechas» en la Figura 9-4:

$$a \to y \to t \qquad b \to x \to s \qquad c \to y \to t$$

b) De acuerdo con la Figura 9-4, los valores imagen bajo f son x e y, y los valores imagen bajo g son r, s y t; por tanto,

$$\text{Im}\, f = \{x, y\} \qquad \text{y} \qquad \text{Im}\, g = \{r, s, t\}$$

Asimismo, por *a*), los valores imagen bajo la aplicación compuesta $g \circ f$ son s y t. En consecuencia, $\text{Im}\,(g \circ f) = \{s, t\}$. Nótese que las imágenes de g y $g \circ f$ son diferentes.

9.3. Considérese la aplicación $F: \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (yz, x^2)$. Calcular:

a) $F(2, 3, 4)$.

b) $F(5, -2, 7)$.

c) $F^{-1}(0, 0)$, es decir, todos los vectores $v \in \mathbf{R}^3$ tales que $F(v) = 0$.

a) Sustituimos en la fórmula para F consiguiendo $F(2, 3, 4) = (3 \cdot 4, 2^2) = (12, 4)$.

b) $F(5, -2, 7) = (-2 \cdot 7, 5^2) = (-14, 25)$.

c) Hacemos $F(v) = 0$ con $v = (x, y, z)$ y despejamos x, y, z:

$$F(x, y, z) = (yz, x^2) = (0, 0) \qquad \text{o} \qquad yz = 0 \quad \text{y} \quad x^2 = 0$$

Así $x = 0$ y bien $y = 0$ o bien $z = 0$. En otras palabras, $x = 0$, $y = 0$ o $x = 0$, $z = 0$. Consecuentemente, v yace bien en el eje z o bien en el eje y.

9.4. Considérese la aplicación $G: \mathbf{R}^3 \to \mathbf{R}^2$ definida por $G(x, y, z) = (x + 2y - 4z, 2x + 3y + z)$. Hallar $G^{-1}(3\ 4)$.

Hacemos $G(x, y, z) = (3, 4)$ para obtener el sistema

$$\begin{aligned} x + 2y - 4z &= 3 \\ 2x + 3y + z &= 4 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y - 4z &= 3 \\ -y + 9z &= -2 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y - 4z &= 3 \\ y - 9z &= 2 \end{aligned}$$

Aquí z es una variable libre. Tomamos $z = a$, $a \in \mathbf{R}$, para llegar a la solución general

$$x = -14a - 1 \qquad y = 9a + 2 \qquad z = a$$

Dicho de otro modo, $G^{-1}(3, 4) = \{(-14a - 1, 9a + 2, a)\}$.

9.5. Considérese la aplicación $F: \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (3y, 2x)$. Sea S el círculo unidad en $\mathbf{R}^2$, esto es, el conjunto solución de $x^2 + y^2 = 1$. *a)* Describir $F(S)$. *b)* Hallar $F^{-1}(S)$.

a) Sea (a, b) un elemento de $F(S)$. Existe un $(x, y) \in S$ tal que $F(x, y) = (a, b)$. Por tanto:

$$(3y, 2x) = (a, b) \qquad \text{o} \qquad 3y = a, 2x = b \qquad \text{o} \qquad y = a/3, x = b/2$$

Como $(x, y) \in S$, es decir, $x^2 + y^2 = 1$, tenemos

$$(b/2)^2 + (a/3)^2 = 1 \qquad \text{o} \qquad a^2/9 + b^2/4 = 1$$

Así pues, $F(S)$ es una elipse.

b) Sea $F(x, y) = (a, b)$, donde $(a, b) \in S$. Entonces $(3y, 2x) = (a, b)$ o $3y = a$, $2x = b$. Dado que $(a, b) \in S$, tenemos que $a^2 + b^2 = 1$. De este modo, $(3y)^2 + (2x)^2 = 1$. En consecuencia, $F^{-1}(S)$ es la elipse $4x^2 + 9y^2 = 1$.

9.6. Defínanse las aplicaciones f y g según $f(x) = 2x + 1$ y $g(x) = x^2 - 2$. Encontrar fórmulas para las aplicaciones *a)* $g \circ f$, *b)* $f \circ g$, *c)* $g \circ g$ (a veces denotada por g^2).

a) Calculamos la fórmula para $g \circ f$ como sigue:

$$(g \circ f)(x) = g(f(x)) = g(2x + 1) = (2x + 1)^2 - 2 = 4x^2 + 4x - 1$$

Obsérvese que puede llegarse a la misma respuesta escribiendo

$$y = f(x) = 2x + 1 \qquad \text{y} \qquad z = g(y) = y^2 - 2$$

y después despejando y como se hace a continuación: $z = y^2 - 2 = (2x+1)^2 - 2 = 4x^2 + 4x - 1$.

b) $(f \circ g)(x) = f(g(x)) = f(x^2 - 2) = 2(x^2 - 2) + 1 = 2x^2 - 3$.

c) $(g \circ g)(x) = g(g(x)) = g(x^2 - 2) = (x^2 - 2)^2 - 2 = x^4 - 4x^2 + 2$.

9.7. Supónganse $f : A \to B$ y $g : B \to C$; existe, pues, la composición $(g \circ f) : A \to C$. Demostrar:

a) Si f y g son inyectivas, $g \circ f$ es inyectiva.

b) Si f y g son suprayectivas, $g \circ f$ es suprayectiva.

c) Si $g \circ f$ es inyectiva, f es inyectiva.

d) Si $g \circ f$ es suprayectiva, g es suprayectiva.

a) Supongamos $(g \circ f)(x) = (g \circ f)(y)$. Entonces $g(f(x)) = g(f(y))$. Como g es inyectiva, $f(x) = f(y)$. Como f es inyectiva, $x = y$. Hemos demostrado que $(g \circ f)(x) = (g \circ f)(y)$ implica $x = y$, luego $g \circ f$ es inyectiva.

b) Supongamos que $c \in C$. Dado que g es suprayectiva, existe $b \in B$ para el que $g(b) = c$. Siendo f suprayectiva, existe $a \in A$ para el que $f(a) = b$. Así $(g \circ f)(a) = g(f(a)) = g(b) = c$; luego $g \circ f$ es suprayectiva.

c) Supongamos que f no es inyectiva. En tal caso, existen elementos distintos x, $y \in A$ tales que $f(x) = f(y)$. De este modo, $(g \circ f)(x) = g(f(x)) = g(f(y)) = (g \circ f)(y)$, luego $g \circ f$ no es inyectiva. Por tanto, si $g \circ f$ es inyectiva, f debe serlo también.

d) Si $a \in A$, entonces $(g \circ f)(a) = g(f(a)) \in g(B)$, luego $(g \circ f)(A) \subseteq g(B)$. Supongamos que g no es suprayectiva. Entonces $g(B)$ estará contenido propiamente en C y por consiguiente $(g \circ f)(A)$ también lo estará, de donde $g \circ f$ no es suprayectiva, En consecuencia si $g \circ f$ es suprayectiva, necesariamente lo es g.

9.8. Demostrar que una aplicación $f : A \to B$ tiene una inversa si y sólo si es inyectiva y suprayectiva.

Supongamos que f tiene una función inversa $f^{-1} : B \to A$ y por tanto $f^{-1} \circ f = 1_A$ y $f \circ f^{-1} = 1_B$. Por ser 1_A inyectiva, f es inyectiva, según el Problema 9.7 *c*), y como 1_B es suprayectiva, f es suprayectiva por el Problema 9.7 *d*). Esto es, f es simultáneamente inyectiva y suprayectiva.

Supongamos ahora que f es tanto inyectiva como suprayectiva. En ese caso, cada $b \in B$ es la imagen de un único elemento de A, digamos $\hat{b}$. De este modo, si $f(a) = b$, necesariamente $a = \hat{b}$; luego $f(\hat{b}) = b$. Denotemos por g la aplicación de B en A definida por $b \mapsto \hat{b}$. Tenemos:

i) $(g \circ f)(a) = g(f(a)) = g(b) = \hat{b} = a$, para todo $a \in A$; por consiguiente, $g \circ f = 1_A$.

ii) $(f \circ g)(b) = f(g(b)) = f(\hat{b}) = b$, para todo $b \in B$; por consiguiente, $f \circ g = 1_B$.

En consecuencia, f tiene por inversa a la aplicación g.

APLICACIONES LINEALES

9.9. Discutir si la siguiente aplicación es lineal: $F : \mathbf{R}^3 \to \mathbf{R}$ definida por $F(x, y, z) = 2x - 3y + 4z$.

Sean $v = (a, b, c)$ y $w = (a', b', c')$; entonces

$$v + w = (a + a', b + b', c + c') \qquad \text{y} \qquad kv = (ka, kb, kc) \qquad k \in \mathbf{R}$$

Tenemos $F(v) = 2a - 3b + 4c$ y $F(w) = 2a' - 3b' + 4c'$. Así

$$F(v + w) = F(a + a', b + b', c + c') = 2(a + a') - 3(b + b') + 4(c + c') =$$
$$= (2a - 3b + 4c) + (2a' - 3b' + 4c') = F(v) + F(w)$$

y
$$F(kv) = F(ka, kb, kc) = 2ka - 3kb + 4kc = k(2a - 3b + 4c) = kF(v)$$

De acuerdo con ello, F es lineal.

9.10. Probar que la siguiente aplicación no es lineal: $F: \mathbf{R}^2 \to \mathbf{R}^3$ definida por

$$F(x, y) = (x + 1, 2y, x + y)$$

Dado que $F(0, 0) = (1, 0, 0) \neq (0, 0, 0)$, F no puede ser lineal.

9.11. Sean V el espacio vectorial de las matrices n-cuadradas sobre K y M una matriz arbitraria en V. Defínase $T: V \to V$ mediante $T(A) = AM + MA$, con $A \in V$. Mostrar que T es lineal.

Para todo par de matrices $A, B \in V$ y todo $k \in K$ tenemos

$$T(A + B) = (A + B)M + M(A + B) = AM + BM + MA + MB =$$
$$= (AM + MA) + (BM + MB) = T(A) + T(B)$$

y
$$T(kA) = (kA)M + M(kA) = k(AM) + k(MA) = k(AM + MA) = kT(A)$$

En consecuencia, T es lineal.

9.12. Demostrar el Teorema 9.2.

La demostración consta de tres pasos: 1. Definimos la aplicación $F: V \to U$ tal que $F(v_i) = u_i$, $i = 1, \ldots, n$. 2. Probamos que F es lineal. 3. Probamos que F es única.

Paso 1. Sea $v \in V$. Como $\{v_1, \ldots, v_n\}$ es una base de V, existen escalares únicos $a_1, \ldots, a_n \in K$ para los que $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$. Definimos $F: V \to U$ por

$$F(v) = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

(Siendo únicos los a_i, la aplicación F está bien definida.) Ahora, para $i = 1, \ldots, n$,

$$v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_n$$

De aquí
$$F(v_i) = 0u_1 + \cdots + 1u_i + \cdots + 0u_n = u_i$$

Así el primer paso de la demostración está completo.

Paso 2. Supongamos $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$ y $w = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n$. Entonces

$$v + w = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \cdots + (a_n + b_n)v_n$$

y, para cualquier $k \in K$, $kv = ka_1 v_1 + ka_2 v_2 + \cdots + ka_n v_n$. Por definición de la aplicación F,

$$F(v) = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n \quad \text{y} \quad F(w) = b_1 u_1 + b_2 v_2 + \cdots + b_n v_n$$

Por tanto,
$$F(v + w) = (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_n + b_n)u_n =$$
$$= (a_1 u_1 + a_2 u_2 + \cdots + a_n u_n) + (b_1 u_1 + b_2 u_2 + \cdots + b_n u_n) =$$
$$= F(v) + F(w)$$

y
$$F(kv) = k(a_1 u_1 + a_2 u_2 + \cdots + a_n u_n) = kF(v)$$

De este modo, F es lineal.

Paso 3. Supongamos que $G: V \to U$ es lineal y que $G(v_i) = u_i$, $i = 1, \ldots, n$. Si

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

necesariamente $G(v) = G(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 G(v_1) + a_2 G(v_2) + \cdots + a_n G(v_n) =$
$= a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = F(v)$

Al ser $G(v) = F(v)$ para todo $v \in V$, $G = F$. Así pues, F es única y queda demostrado el teorema.

9.13. Sea $F: \mathbf{R}^2 \to \mathbf{R}^2$ la aplicación lineal para la cual

$$F(1, 2) = (2, 3) \qquad \text{y} \qquad F(0, 1) = (1, 4)$$

[Dado que (1, 2) y (0, 1) constituyen una base de $\mathbf{R}^2$, tal aplicación lineal existe y es única en virtud del Teorema 9.2.] Encontrar una expresión para F, es decir, hallar $F(a, b)$.

Escribimos (a, b) como combinación lineal de (1, 2) y (0, 1) usando incógnitas x e y:

$$(a, b) = x(1, 2) + y(0, 1) = (x, 2x + y) \qquad \text{de forma que} \qquad a = x,\ b = 2x + y$$

Despejamos x e y en términos de a y b llegando a $x = a$, $y = -2a + b$. Entonces

$$F(a, b) = xF(1, 2) + yF(0, 1) = a(2, 3) + (-2a + b)(1, 4) = (b, -5a + 4b)$$

9.14. Sea $T: V \to U$ lineal y supóngase que $v_1, \ldots, v_n \in V$ poseen la propiedad de que sus imágenes $T(v_1), \ldots, T(v_n)$ son linealmente independientes. Mostrar que los vectores $v_1, \ldots, v_n$ son también linealmente independientes.

Supongamos que, para ciertos escalares $a_1, \ldots, a_n$, $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$. En tal caso,

$$0 = T(0) = T(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 T(v_1) + a_2 T(v_2) + \cdots + a_n T(v_n)$$

Como los $T(v_i)$ son linealmente independientes, todo los $a_i = 0$. Los vectores $v_1, \ldots, v_n$ son, pues, linealmente independientes.

9.15. Supóngase que la aplicación lineal $F: V \to U$ es inyectiva y suprayectiva. Probar que la aplicación inversa $F^{-1}: U \to V$ es también lineal.

Supongamos $u, u' \in U$. Por ser F inyectiva y suprayectiva, existen vectores únicos $v, v' \in V$ para los que $F(v) = u$ y $F(v') = u'$. Como F es lineal, tendremos también

$$F(v + v') = F(v) + F(v') = u + u' \qquad \text{y} \qquad F(kv) = kF(v) = ku$$

Por definición de aplicación inversa, $F^{-1}(u) = v$, $F^{-1}(u') = v'$, $F^{-1}(u + u') = v + v'$ y $F^{-1}(ku) = kv$. Entonces

$$F^{-1}(u + u') = v + u' = F^{-1}(u) + F^{-1}(u') \qquad \text{y} \qquad F^{-1}(ku) = kv = kF^{-1}(u)$$

y así F^{-1} es lineal.

IMAGEN Y NUCLEO DE APLICACIONES LINEALES

9.16. Sea $F: \mathbf{R}^5 \to \mathbf{R}^3$ la aplicación lineal definida según

$$F(x, y, z, s, t) = (x + 2y + z - 3s + 4t,\ 2x + 5y + 4z - 5s + 5t,\ x + 4y + 5z - s - 2t)$$

Hallar una base y la dimensión de la imagen de F.

Determinamos la imagen de los vectores de la base usual de $\mathbf{R}^5$:

$$F(1, 0, 0, 0, 0) = (1, 2, 1) \qquad F(0, 1, 0, 0, 0) = (2, 5, 4) \qquad F(0, 0, 1, 0, 0) = (1, 4, 5)$$
$$F(0, 0, 0, 1, 0) = (-3, -5, -1) \qquad F(0, 0, 0, 0, 1) = (4, 5, -2)$$

De acuerdo con la Proposición 9.4, los vectores imagen generan Im F, por lo que construimos la matriz cuyas filas son estos vectores imagen y la reducimos por filas a forma escalonada:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 4 \\ 1 & 4 & 5 \\ -3 & -5 & -1 \\ 4 & 5 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Así pues, $(1, 2, 1)$ y $(0, 1, 2)$ forman una base de Im F, de modo que dim (Im F) = 2.

9.17. Sea $G: \mathbf{R}^3 \to \mathbf{R}^3$ la aplicación lineal definida según

$$G(x, y, z) = (x + 2y - z, y + z, x + y - 2z)$$

Hallar una base y la dimensión del núcleo de G.

Hacemos $G(v) = 0$, donde $v = (x, y, z)$:

$$G(x, y, z) = (x + 2y - z, y + z, x + y - 2z) = (0, 0, 0)$$

Igualamos entre sí las componentes correspondientes para formar el sistema homogéneo cuyo espacio solución es el núcleo W de G:

$$\begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ x + y - 2z &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ -y - z &= 0 \end{aligned} \qquad \text{o} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \end{aligned}$$

La única variable libres es z; por tanto, dim $W = 1$. Sea $z = 1$; entonces $y = -1$ y $x = 3$. Así $(3, -1, 1)$ forma una base de Ker G.

9.18. Considérese la aplicación matricial A: $\mathbf{R}^4 \to \mathbf{R}^3$, donde $A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 3 & 5 & -2 \\ 3 & 8 & 13 & -3 \end{pmatrix}$. Hallar una base y la dimensión de *a*) la imagen de A, *b*) el núcleo de A.

a) El espacio columna de A es igual a Im A, por lo que reducimos A^T a forma escalonada:

$$A^T = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 8 \\ 3 & 5 & 13 \\ 1 & -2 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 0 & -3 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

De este modo, $\{(1, 1, 3), (0, 1, 2)\}$ es una base de Im A y dim (Im A) = 2.

b) Aquí Ker A es el espacio solución del sistema homogéneo $AX = 0$, donde $X = (x, y, z, t)^T$. Por eso reducimos la matriz A de coeficientes a forma escalonada:

$$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -3 \\ 0 & 2 & 4 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{cases} x + 2y + 3z + \ \ t = 0 \\ \qquad y + 2z - 3t = 0 \end{cases}$$

Las variables libres son z y t, luego dim (Ker A) = 2. Tomemos:

i) $z = 1$, $t = 0$, para conseguir la solución $(1, -2, 1, 0)$.

ii) $z = 0$, $t = 1$, para conseguir la solución $(-7, 3, 0, 1)$.

Así pues, $(1, -2, 1, 0)$ y $(-7, 3, 0, 1)$ constituyen una base de Ker A.

9.19. Considérese la aplicación matricial $B: \mathbf{R}^3 \to \mathbf{R}^3$ con $B = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 5 & 13 \\ -2 & -1 & -4 \end{pmatrix}$. Encontrar la dimensión y una base de *a*) el núcleo de B, *b*) la imagen de B.

a) Reducimos B a forma escalonada para conseguir el sistema homogéneo correspondiente a Ker B:

$$B = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 5 & 13 \\ -2 & -1 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 5 \\ 0 & -1 & -2 \\ 0 & 3 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{cases} x + 2y + 5z = 0 \\ \qquad y + 2z = 0 \end{cases}$$

Hay una variable libre, z, luego dim (Ker B) = 1. Tomamos $z = 1$ para llegar a la solución $(-1, -2, 1)$ que forma una base de Ker B.

b) Reducimos B^T a forma escalonada:

$$B^T = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 5 & -1 \\ 5 & 13 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 \\ 0 & -1 & 3 \\ 0 & -2 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Así $(1, 3, -2)$ y $(0, 1, -3)$ constituyen una base de Im B.

9.20. Encontrar una aplicación lineal $F: \mathbf{R}^3 \to \mathbf{R}^4$ cuya imagen esté generada por $(1, 2, 0, -4)$ y $(2, 0, -1, -3)$.

Método 1. Consideremos la base usual de $\mathbf{R}^3$: $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Tomemos

$$F(e_1) = (1, 2, 0, -4) \qquad F(e_2) = (2, 0, -1, -3) \qquad \text{y} \qquad F(e_3) = (0, 0, 0, 0)$$

Según el Teorema 9.2, tal aplicación lineal F existe y es única. Además, la imagen de F está generada por los $F(e_i)$. Por consiguiente, F posee la propiedad requerida. Hallemos una fórmula para $F(x, y, z)$:

$$\begin{aligned} F(x, y, z) &= F(xe_1 + ye_2 + ze_3) = xF(e_1) + yF(e_1) + zF(e_3) = \\ &= x(1, 2, 0, -4) + y(2, 0, -1, -3) + z(0, 0, 0, 0) = \\ &= (x + 2y, 2x, -y, -4x - 3y) \end{aligned}$$

Método 2. Construyamos una matriz 4×3 A cuyas columnas consistan sólo en los vectores dados; es decir,

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 0 \\ 0 & -1 & -1 \\ -4 & -3 & -3 \end{pmatrix}$$

Recuérdese que A determina una aplicación lineal $A: \mathbf{R}^3 \to \mathbf{R}^4$ con imagen generada por las columnas de A. Siendo así, A satisface la condición requerida.

9.21. Sean V el espacio vectorial de las matrices 2 por 2 sobre $\mathbf{R}$ y $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Sea $F: V \to V$ la aplicación lineal definida por $F(A) = AM - MA$. Hallar una base y la dimensión del núcleo W de F.

Buscamos el conjunto de las $\begin{pmatrix} x & y \\ s & t \end{pmatrix}$ tales que $F\begin{pmatrix} x & y \\ s & t \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

$$F\begin{pmatrix} x & y \\ s & t \end{pmatrix} = \begin{pmatrix} x & y \\ s & t \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ s & t \end{pmatrix} =$$

$$= \begin{pmatrix} x & 2x + 3y \\ s & 2s + 3t \end{pmatrix} - \begin{pmatrix} x + 2s & y + 2t \\ 3s & 3t \end{pmatrix} =$$

$$= \begin{pmatrix} -2s & 2x + 2y - 2t \\ -2s & 2s \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Por tanto,

$$\begin{cases} 2x + 2y - 2t = 0 \\ 2s = 0 \end{cases} \quad \text{o} \quad \begin{cases} x + y - t = 0 \\ s = 0 \end{cases}$$

Las variables libres son y y t, luego $\dim W = 2$. Para conseguir una base de W tomemos

a) $y = -1$, $t = 0$, para obtener la solución $x = 1$, $y = -1$, $s = 0$, $t = 0$.

b) $y = 0$, $t = 1$, para obtener la solución $x = 1$, $y = 0$, $s = 0$, $t = 1$.

Así pues, $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ es una base de W.

9.22. Demostrar el Teorema 9.3.

a) Como $F(0) = 0$, tenemos que $0 \in \operatorname{Im} F$. Supongamos ahora que u, $u' \in \operatorname{Im} F$ y a, $b \in K$. Dado que u y u' pertenecen a la imagen de F, existen vectores v, $v' \in V$ tales que $F(v) = u$ y $F(v') = u'$. Entonces

$$F(av + bv') = aF(v) + bF(v') = au + bu' \in \operatorname{Im} F$$

De este modo, la imagen de F es un subespacio de U.

b) Como $F(0) = 0$, tenemos que $0 \in \operatorname{Ker} F$. Supongamos ahora que v, $w \in \operatorname{Ker} F$ y a, $b \in K$. Dado que v y w pertenecen al núcleo de F, $F(v) = 0$ y $F(w) = 0$. Así

$$F(av + bw) = aF(v) + bF(w) = a0 + b0 = 0 \qquad \text{luego} \qquad av + bw \in \operatorname{Ker} F$$

El núcleo de F es, pues, un subespacio de V.

9.23. Demostrar el Teorema 9.5.

Supongamos que dim (Ker F) $= r$ y que $\{w_1, \ldots, w_r\}$ es una base de Ker F. Supongamos, asimismo, que dim (Im F) $= s$ y que $\{u_1, \ldots, u_s\}$ es una base de Im F. (Por la Proposición 9.4 sabemos que Im F tiene dimensión finita.) Dado que los $u_j \in \text{Im } F$, existen vectores $v_1, \ldots, v_s$ en V tales que $F(v_1) = u_1, \ldots, F(v_s) = u_s$. Afirmamos que el conjunto

$$B = \{w_1, \ldots, w_r, v_1, \ldots, v_s\}$$

es una base de V, esto es: i) B genera V, ii) B es linealmente independiente. Una vez que hayamos probado i) y ii), tendremos dim $V = r + s =$ dim (Ker F) + dim (Im F).

i) *B genera V.*
Sea $v \in V$. En tal caso, $F(v) \in \text{Im } F$. Como los u_j generan Im F, existen escalares $a_1, \ldots, a_s$ tales que $F(v) = a_1 u_1 + \cdots + a_s u_s$. Tomemos $\hat{v} = a_1 v_1 + \cdots + a_s v_s - v$. Entonces

$$F(\hat{v}) = F(a_1 v_1 + \cdots + a_s v_s - v) = a_1 F(v_1) + \cdots + a_s F(v_s) - F(v) =$$
$$= a_1 u_1 + \cdots + a_s u_s - F(v) = 0$$

Por tanto, $\hat{v} \in \text{Ker } F$. Como los w_i generan Ker F, existen escalares $b_1, \ldots, b_r$ tales que

$$\hat{v} = b_1 w_1 + \cdots + b_r w_r = a_1 v_1 + \cdots + a_s v_s - v$$

En consecuencia,

$$v = a_1 v_1 + \cdots + a_s v_s - b_1 w_1 - \cdots - b_r w_r$$

Siendo así, B genera V.

ii) *B es linealmente independiente.*
Supongamos que

$$x_1 w_1 + \cdots + x_r w_r + y_1 v_1 + \cdots + y_s v_s = 0 \qquad [1]$$

donde $x_i, y_j \in K$. Entonces

$$0 = F(0) = F(x_1 w_1 + \cdots + x_r w_r + y_1 v_1 + \cdots + y_s v_s) =$$
$$= x_1 F(w_1) + \cdots + x_r F(w_r) + y_1 F(v_1) + \cdots + y_s F(v_s) \qquad [2]$$

Pero $F(w_i) = 0$ porque $w_i \in \text{Ker } F$ y $F(v_j) = u_j$. La sustitución en [2] proporciona $y_1 u_1 + \cdots + y_s u_s = 0$. Por ser linealmente independientes los u_j, cada $y_j = 0$. La sustitución en [1] conduce a $x_1 w_1 + \cdots + x_r w_r = 0$. Siendo linealmente independientes los w_i, cada $x_i = 0$. De modo que B es linealmente independiente.

9.24. Supóngase que $F: V \to U$ y $G: U \to W$ son lineales. Demostrar:

a) rango $(G \circ F) \leq$ rango G. *b*) rango $(G \circ F) \leq$ rango F.

a) Como $F(V) \subseteq U$, tendremos también $G(F(V)) \subseteq G(U)$ y así dim $G(F(V)) \leq$ dim $G(U)$. Entonces

$$\text{rango } (G \circ F) = \dim ((G \circ F)(V)) = \dim (G(F(V))) \leq \dim G(U) = \text{rango } G$$

b) Tenemos dim $(G(F(V))) \leq$ dim $F(V)$. De aquí

$$\text{rango } (G \circ F) = \dim ((G \circ F)(V)) = \dim (G(F(V))) \leq \dim F(V) = \text{rango } F$$

9.25. Supóngase que $f: V \to U$ es lineal con núcleo W y que $f(v) = u$. Mostrar que la «variedad afín» $v + W = \{v + w : w \in W\}$ es la preimagen de u, es decir, $f^{-1}(u) = v + W$.

Debemos demostrar que: i) $f^{-1}(u) \subseteq v + W$, ii) $v + W \subseteq f^{-1}(u)$. Comenzamos por i). Supongamos que $v' \in f^{-1}(u)$. Entonces $f(v') = u$, de forma que

$$f(v' - v) = f(v') - f(v) = u - u = 0$$

esto es, $v' - v \in W$. Así $v' = v + (v' - v) \in v + W$ y por tanto $f^{-1}(u) \subseteq v + W$.

Demostremos ahora ii). Supongamos que $v' \in v + W$. Entonces $v' = v + w$, donde $w \in W$. Dado que W es el núcleo de f, $f(w) = 0$. En consecuencia,

$$f(v') = f(v + w) + f(v) + f(w) = f(v) + 0 = f(v) = u$$

Así $v' \in f^{-1}(u)$, luego $v + W \subseteq f^{-1}(u)$.

APLICACIONES LINEALES SINGULARES Y NO SINGULARES. ISOMORFISMOS

9.26. Determinar si cada aplicación lineal es o no singular. Si es singular, encontrar un vector no nulo v cuya imagen sea 0.

a) $F: \mathbf{R}^2 \to \mathbf{R}^2$ definida por $F(x, y) = (x - y, x - 2y)$.

b) $G: \mathbf{R}^2 \to \mathbf{R}^2$ definida por $G(x, y) = (2x - 4y, 3x - 6y)$.

a) Hallamos Ker F haciendo $F(v) = 0$, donde $v = (x, y)$:

$$(x - y, x - 2y) = (0, 0) \quad \text{o} \quad \begin{cases} x - y = 0 \\ x - 2y = 0 \end{cases} \quad \text{o} \quad \begin{cases} x - y = 0 \\ -y = 0 \end{cases}$$

La única solución es $x = 0$, $y = 0$, luego F es no singular.

b) Hacemos $G(x, y) = (0, 0)$ para hallar Ker G:

$$(2x - 4y, 3x - 6y) = (0, 0) \quad \text{o} \quad \begin{cases} 2x - 4y = 0 \\ 3x - 6y = 0 \end{cases} \quad \text{o} \quad x - 2y = 0$$

El sistema tiene soluciones no triviales porque y es una variable libre; por tanto, G es singular. Tomemos $y = 1$ para obtener la solución $v = (2, 1)$, que es un vector no nulo tal que $G(v) = 0$.

9.27. Defínase $H: \mathbf{R}^3 \to \mathbf{R}^3$ por $H(x, y, z) = (x + y - 2z, x + 2y + z, 2x + 2y - 3z)$. *a*) Probar que H es no singular. *b*) Hallar una expresión para H^{-1}.

a) Hacemos $H(x, y, z) = (0, 0, 0)$; es decir, hacemos

$$(x + y - 2z, x + 2y + z, 2x + 2y - 3z) = (0, 0, 0)$$

Esto conduce al sistema homogéneo

$$\begin{cases} x + y - 2z = 0 \\ x + 2y + z = 0 \\ 2x + 2y - 3z = 0 \end{cases} \quad \text{o} \quad \begin{cases} x + y - 2z = 0 \\ \phantom{x +{}} y + 3z = 0 \\ \phantom{x + y + 3{}} z = 0 \end{cases}$$

El sistema escalonado está en forma triangular, por lo que la única solución es $x = 0$, $y = 0$, $z = 0$. Siendo así, H es no singular.

b) Hacemos $H(x, y, z) = (a, b, c)$ y despejamos x, y, z en términos de a, b, c:

$$\begin{cases} x + y - 2z = a \\ x + 2y + z = b \\ 2x + 2y - 3z = c \end{cases} \quad \text{o} \quad \begin{cases} x + y - 2z = a \\ \phantom{x +{}} y + 3z = b - a \\ \phantom{x + y + 3{}} z = c - 2a \end{cases}$$

Resolviendo para x, y, z, $x = -8a - b + 5c$, $y = 5a + b - 3c$, $z = -2a + c$. De este modo,

$$H^{-1}(a, b, c) = (-8a - b + 5c, 5a + b - 3c, -2a + c)$$

o sustituyendo a, b, c por x, y, z, respectivamente,

$$H^{-1}(x, y, z) = (-8x - y + 5z, 5x + y - 3z, -2x + z)$$

9.28. Supóngase que $F: V \to U$ es lineal y que V es de dimensión finita. Probar que V y la imagen de F tienen la misma dimensión si y sólo si F es no singular. Determinar todas las aplicaciones lineales no singulares $T: \mathbf{R}^4 \to \mathbf{R}^3$.

De acuerdo con el Teorema 9.5, dim V = dim (Im F) + dim (Ker F). Por consiguiente, V e Im F tienen la misma dimensión si y sólo si dim (Ker F) = 0 o Ker $F = \{0\}$, o sea, si y sólo si F es no singular.

Como dim $\mathbf{R}^3$ es menor que dim $\mathbf{R}^4$, tendremos que dim (Im T) es menor que la dimensión del dominio $\mathbf{R}^4$ de T. En consecuencia, ninguna aplicación lineal $T: \mathbf{R}^4 \to \mathbf{R}^3$ puede ser no singular.

9.29. Demostrar el Teorema 9.7.

Supongamos que $v_1, v_2, \ldots, v_n$ son vectores linealmente independientes en V. Afirmamos que $F(v_1), F(v_2), \ldots, F(v_n)$ son también linealmente independientes. Supongamos $a_1 F(v_1) + a_2 F(v_2) + \cdots + a_n F(v_n) = 0$, donde $a_i \in K$. Dado que F es lineal, $F(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = 0$; por tanto,

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \in \text{Ker } F$$

Pero F es no singular, esto es, Ker $F = \{0\}$, luego $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$. Siendo linealmente independientes los v_i, todos los a_i son 0. De acuerdo con esto, los $F(v_i)$ son linealmente independientes. Queda así demostrado el teorema.

9.30. Demostrar el Teorema 9.9.

Si F es un isomorfismo, sólo 0 se aplica en 0, de modo que F es no singular. Supongamos que F es no singular. Entonces dim (Ker F) = 0. Según el Teorema 9.5, dim V = dim (Ker F) + dim (Im F). Así pues, dim U = dim V = dim (Im F). Como U tiene dimensión finita, Im $F = U$ y por ende F es suprayectiva. Por tanto, F es simultáneamente inyectiva y suprayectiva, es decir, F es un isomorfismo.

OPERACIONES CON APLICACIONES LINEALES

9.31. Defínanse $F: \mathbf{R}^3 \to \mathbf{R}^2$ y $G: \mathbf{R}^3 \to \mathbf{R}^2$ por $F(x, y, z) = (2x, y + z)$ y $G(x, y, z) = (x - z, y)$, respectivamente. Encontrar expresiones que definan las aplicaciones *a*) $F + G$, *b*) $3F$, *c*) $2F - 5G$.

a) $(F + G)(x, y, z) = F(x, y, z) + (G(x, y, z) =$

$$= (2x, y + z) + (x - z, y) = (3x - z, 2y + z).$$

b) $(3F)(x, y, z) = 3F(x, y, z) = 3(2x, y + z) = (6x, 3y + 3z).$

c) $(2F - 5G)(x, y, z) = 2F(x, y, z) - 5G(x, y, z) = 2(2x, y + z) - 5(x - z, y) =$

$$= (4x, 2y + 2z) + (-5x + 5z, -5y) = (-x + 5z, -3y + 2z).$$

9.32. Defínanse $F:\mathbf{R}^3 \to \mathbf{R}^2$ y $G:\mathbf{R}^2 \to \mathbf{R}^2$ por $F(x, y, z) = (2x, y + z)$ y $G(x, y) = (y, x)$, respectivamente. Derivar fórmulas que definan las aplicaciones *a*) $G \circ F$, *b*) $F \circ G$.

a) $(G \circ F)(x, y, z) = G(F(x, y, z)) = G(2x, y + z) = (y + z, 2x)$.

b) La aplicación $F \circ G$ no está definida, ya que la imagen de G no está contenida en el dominio de F.

9.33. Demostrar las siguientes aserciones: *a*) La aplicación cero, $\mathbf{0}$, definida según $\mathbf{0}(v) = 0$ para todo $v \in V$ es el elemento cero de Hom (V, U). *b*) El opuesto de $F \in \text{Hom}\,(V, U)$ es la aplicación $(-1)F$, o sea, $-F = (-1)F$.

a) Sea $F \in \text{Hom}\,(V, U)$. Para todo $v \in V$,

$$(F + \mathbf{0})(v) = F(v) + \mathbf{0}(v) = F(v) + 0 = F(v)$$

Como $(F + \mathbf{0})(v) = F(v)$ para todo $v \in V$, $F + \mathbf{0} = F$.

b) Para todo $v \in V$,

$$(F + (-1)F)(v) = F(v) + (-1)F(v) = F(v) - F(v) = 0 = \mathbf{0}(v)$$

Siendo $(F + (-1)F)(v) = \mathbf{0}(v)$ para todo $v \in V$, $F + (-1)F = \mathbf{0}$. De esta manera, $(-1)F$ es el opuesto de F.

9.34. Supóngase que $F_1, F_2, \ldots, F_n$ son aplicaciones lineales de V en U. Mostrar que para escalares cualesquiera $a_1, a_2, \ldots, a_n$ y para todo $v \in V$,

$$(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) = a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)$$

Por definición de la aplicación a_1F_1, $(a_1F_1)(v) = a_1F_1(v)$, luego el teorema es válido para $n = 1$. Por inducción,

$$\begin{aligned}(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) &= (a_1F_1)(v) + (a_2F_2 + \cdots + a_nF_n)(v) = \\ &= a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)\end{aligned}$$

9.35. Considérense las aplicaciones lineales $F:\mathbf{R}^3 \to \mathbf{R}^2$, $G:\mathbf{R}^3 \to \mathbf{R}^2$, $H:\mathbf{R}^3 \to \mathbf{R}^2$ definidas por

$$F(x, y, z) = (x + y + z, x + y) \qquad G(x, y, z) = (2x + z, x + t) \qquad H(x, y, z) = (2y, x)$$

Probar que F, G, H son linealmente independientes [como elementos de Hom (V, U)].

Supongamos que, para ciertos escalares $a, b, c \in K$,

$$aF + bG + cH = \mathbf{0} \tag{1}$$

(Aquí $\mathbf{0}$ es la aplicación cero.) Para $e_1 = (1, 0, 0) \in \mathbf{R}^3$ tenemos

$$\begin{aligned}(aF + bG + cH)(e_2) &= aF(0, 1, 0) + bG(0, 1, 0) + cH(0, 1, 0) = \\ &= a(1, 1) + b(0, 1) + c(2, 0) = (a + 2c, a + b) = \mathbf{0}(e_2) = (0, 0)\end{aligned}$$

y $\mathbf{0}(e_1) = (0, 0)$. Así, por [1], $(a + 2b, a + b + c) = (0, 0)$ y por tanto

$$a + 2b = 0 \qquad \text{y} \qquad a + b + c = 0 \tag{2}$$

De forma similar, para $e_2 = (0, 1, 0) \in \mathbf{R}^3$ tenemos

$$\begin{aligned}(aF + bG + cH)(e_2) &= aF(0, 1, 0) + bG(0, 1, 0) + cH(0, 1, 0) = \\ &= a(1, 1) + b(0, 1) + c(2, 0) = (a + 2c, a + b) = \mathbf{0}(e_2) = (0, 0)\end{aligned}$$

De este modo, $a + 2c = 0$ y $a + b = 0$ [3]

Usando [2] y [3] obtenemos $a = 0$ $b = 0$ $c = 0$ [4]

Como [1] implica [4], las aplicaciones F, G y H son linealmente independientes.

9.36. Demostrar el Teorema 9.11.

Supongamos que $\{v_1, \ldots, v_m\}$ es una base de V y que $\{u_1, \ldots, u_n\}$ es una de U. De acuerdo con el Teorema 9.2, una aplicación lineal en Hom (V, U) queda unívocamente determinada asignando arbitrariamente elementos de U a los elementos v_i de la base de V. Definamos

$$F_{ij} \in \text{Hom}\,(V, U) \qquad i = 1, \ldots, m, j = 1, \ldots, n$$

como la aplicación lineal para la cual $F_{ij}(v_i) = u_j$ y $F_{ij}(v_k) = 0$ si $k \neq i$. Esto es, F_{ij} aplica v_i en u_j y el resto de los v en 0. Obsérvese que $\{F_{ij}\}$ contiene exactamente mn elementos. Por consiguiente, el teorema quedará demostrado si probamos que es una base de Hom (V, U).

Demostración de que $\{F_{ij}\}$ *genera* Hom (V, U). Consideremos una función arbitraria $F \in$ Hom (V, U). Supongamos $F(v_1) = w_1$, $F(v_2) = w_2$, $\ldots$, $F(v_m) = w_m$. Dado que $w_k \in U$ es una combinación lineal de los u; es decir,

$$w_k = a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n \qquad k = 1, \ldots, m, \quad a_{ij} \in K \tag{1}$$

Consideremos la aplicación lineal $G = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}$. Siendo G una combinación lineal de los F_{ij}, la demostración de que $\{F_{ij}\}$ genera Hom (V, U) se completará en cuanto probemos que $F = G$.

Calculamos ahora $G(v_k)$, $k = 1, \ldots, m$. Como $F_{ij}(v_k) = 0$ para $k \neq i$ y $F_{ki}(v_k) = u_i$,

$$G(v_k) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j =$$
$$= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n$$

Así, según [1], $G(v_k) = w_k$ para cada k. Pero $F(v_k) = w_k$ para cada k. En consecuencia, y en virtud del Teorema 9.2, $F = G$ y $\{F_{ij}\}$ genera Hom (V, U).

Demostración de que $\{F_{ij}\}$ *es linealmente independiente*. Supongamos, para escalares $a_{ij} \in K$,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij} = \mathbf{0}$$

Para v_k, $k = 1, \ldots, m$,

$$0 = \mathbf{0}(v_k) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j =$$
$$= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n$$

Pero los u_i son linealmente independientes, luego para $k = 1, \ldots, m$ tenemos $a_{k1} = 0$, $a_{k2} = 0$, $\ldots$, $a_{kn} = 0$. Dicho de otro modo, todos los $a_{ij} = 0$ y por tanto $\{F_{ij}\}$ es linealmente independiente.

9.37. Demostrar el Teorema 9.12.

i) Para todo $v \in V$,

$$(G \circ (F + F'))(v) = G((F + F')(v)) = G(F(v) + F'(v)) =$$
$$= G(F(v)) + G(F'(v)) = (G \circ F)(v) + (G \circ F')(v) = (G \circ F + G \circ F')(v)$$

Así $G \circ (F + F') = G \circ F + G \circ F'$.

ii) Para todo $v \in V$,

$$((G + G') \circ F)(v) = (G + G')(F(v)) = G(F(v)) + G'(F(v)) =$$
$$= (G \circ F)(v) + (G' \circ F)(v) = (G \circ F + G' \circ F)(v)$$

Así $(G + G') \circ F = G \circ F + G' \circ F$.

iii) Para todo $v \in V$,

$$(k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = (kG)(F(v)) = (kG \circ F)(v)$$

y $$(k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = G(kF(v)) = G((kF)(v)) = (G \circ kF)(v)$$

En consecuencia, $k(G \circ F) = (kG) \circ F = G \circ (kF)$. (Señalamos que se prueba que dos aplicaciones son iguales mostrando que asignan la misma imagen a cada punto del dominio.)

ALGEBRA DE OPERADORES LINEALES

9.38. Sean S y T los operadores lineales en $\mathbf{R}^2$ definidos por $S(x, y) = (y, x)$ y $T(x, y) = (0, x)$. Encontrar las expresiones que definen los operadores *a*) $S + T$, *b*) $2S - 3T$, *c*) ST, *d*) TS, *e*) S^2, *f*) T^2.

a) $(S + T)(x, y) = S(x, y) + T(x, y) = (y, x) + (0, x) = (y, 2x)$.

b) $(2S - 3T)(x, y) = 2S(x, y) - 3T(x, y) = 2(y, x) - 3(0, x) = (2y, -x)$.

c) $(ST)(x, y) = S(T(x, y)) = S(0, x) = (x, 0)$.

d) $(TS)(x, y) = T(S(x, y)) = T(y, x) = (0, y)$.

e) $S^2(x, y) = S(S(x, y)) = S(y, x) = (x, y)$. Nótese que $S^2 = I$, la aplicación identidad.

f) $T^2(x, y) = T(T(x, y)) = T(0, x) = (0, 0)$. Nótese que $T^2 = \mathbf{0}$, la aplicación cero.

9.39. Considérese el operador lineal T en $\mathbf{R}^3$ definido por $T(x, y, z) = (2x, 4x - y, 2x + 3y - z)$. *a*) Probar que T es invertible. Encontrar fórmulas para: *b*) T^{-1}, *c*) T^2, *d*) T^{-2}.

a) Sea $W = \operatorname{Ker} T$. Sólo necesitamos probar que T es no singular, o sea, que $W = \{0\}$. Hacemos $T(x, y, z) = (0, 0, 0)$, lo que conduce a

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (0, 0, 0)$$

Así W es el espacio solución del sistema homogéneo

$$2x = 0 \qquad 4x - y = 0 \qquad 2x + 3y - z = 0$$

que tiene sólo la solución trivial $(0, 0, 0)$. De este modo, $W = \{0\}$, luego T es no singular y por ende invertible.

b) Hacemos $T(x, y, z) = (r, s, t)$ [y así $T^{-1}(r, s, t) = (x, y, z)$]. Tenemos

$$(2x, 4x - y, 2x + 3y - z) = (r, s, t) \qquad \text{o} \qquad 2x = r,\ 4x - y = s,\ 2x + 3y - z = t$$

Despejamos x, y, z en términos de r, s, t obteniendo $x = \frac{1}{2}r$, $y = 2r - s$, $z = 7r - 3s - t$. Así pues,

$$T^{-1}(r, s, t) = (\tfrac{1}{2}r, 2r - s, 7r - 3s - t) \qquad \text{o} \qquad T^{-1}(x, y, z) = (\tfrac{1}{2}x, 2x - y, 7x - 3y - z)$$

c) Aplicamos T dos veces consiguiendo

$$\begin{aligned} T^2(x, y, z) &= T(2x, 4x - y, 2x + 3y - z) = \\ &= [4x, 4(2x) - (4x - y), 2(2x) + 3(4x - y) - (2x + 3y - z)] = \\ &= (4x, 4x + y, 14x - 6y + z) \end{aligned}$$

d) Aplicamos T^{-1} dos veces consiguiendo

$$\begin{aligned} T^{-2}(x, y, z) &= T^{-2}(\tfrac{1}{2}x, 2x - y, 7x - 3y - z) = \\ &= [\tfrac{1}{4}x, 2(\tfrac{1}{2}x) - (2x - y), 7(\tfrac{1}{2}x) - 3(2x - y) - (7x - 3y - z)] = \\ &= (\tfrac{1}{4}x, -x + y, -\tfrac{19}{2}x + 6y + z) \end{aligned}$$

9.40. Sea V de dimensión finita y T un operador lineal en V tal que $TS = I$ para algún operador S en V. (Llamamos a S una inversa por la derecha de T.) *a*) Probar que T es invertible. *b*) Probar que $S = T^{-1}$. *c*) Dar un ejemplo que muestre que lo anterior no es necesariamente válido si V es de dimensión infinita.

a) Sea dim $V = n$. Según el Teorema 9.14, T es invertible si y sólo si es suprayectivo, de donde T es invertible si y sólo si rango $T = n$. Tenemos $n =$ rango $I =$ rango $TS \leq$ rango $T \leq n$. Por consiguiente, ranto $T = n$ y T es invertible

b) $TT^{-1} = T^{-1}T = I$. Entonces $S = IS = (T^{-1}T)S = T^{-1}(TS) = T^{-1}I = T^{-1}$.

c) Sea V el espacio de los polinomios sobre K; digamos $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$. Sean T y S los operadores en V definidos por

$$T(p(t)) = 0 + a_1 + a_2 t + \cdots + a_n t^{n-1} \qquad \text{y} \qquad S(p(t)) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$$

Tenemos

$$(TS)(p(t)) = T(S(p(t))) = T(a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}) = a_0 + a_1 t + \cdots + a_n t^n = p(t)$$

de modo que $TS = I$, la aplicación identidad. Por otra parte, si $k \in K$ y $k \neq 0$, entonces

$$(ST)(k) = S(T(k)) = S(0) = 0 \neq k$$

En consecuencia, $ST \neq I$.

9.41. Sean S y T operadores lineales en $\mathbf{R}^2$ definidos por $S(x, y) = (0, x)$ y $T(x, y) = (x, 0)$. Probar que $TS = \mathbf{0}$ pero $ST \neq \mathbf{0}$. Probar, asimismo, que $T^2 = T$.

$(TS)(x, y) = T(S(x, y)) = T(0, x) = (0, 0)$. Como TS asigna $0 = (0, 0)$ a todo $(x, y) \in \mathbf{R}^2$, es la aplicación cero: $TS = \mathbf{0}$.

$(ST)(x, y) = S(T(x, y)) = S(x, 0) = (0, x)$. Por ejemplo, $(ST)(4, 2) = (0, 4)$. Siendo así, $ST \neq \mathbf{0}$, porque no asigna $0 = (0, 0)$ a todo elemento de $\mathbf{R}^2$.

Para todo $(x, y) \in \mathbf{R}^2$, $T^2(x, y) = T(T(x, y)) = T(x, 0) = (x, 0) = T(x, y)$. De aquí $T^2 = T$.

9.42. Considérese el operador lineal T en $\mathbf{R}^2$ definido por $T(x, y) = (2x + 4y, 3x + 6y)$. Hallar: *a*) una fórmula para T^{-1}, *b*) $T^{-1}(8, 12)$, *c*) $T^{-1}(1, 2)$. *d*) ¿Es T una aplicación suprayectiva?

a) T es singular; por ejemplo, $T(2, 1) = (0, 0)$. Por consiguiente, el operador lineal $T^{-1}: \mathbf{R}^2 \to \mathbf{R}^2$ no existe.

b) $T^{-1}(8, 12)$ quiere decir la preimagen de $(8, 12)$ bajo T. Hacemos $T(x, y) = (8, 12)$ obteniendo el sistema

$$\begin{cases} 2x + 4y = 8 \\ 3x + 6y = 12 \end{cases} \quad \text{o} \quad x + 2y = 4$$

Aquí y es una variable libre. Tomamos $y = a$, donde a es un parámetro, para conseguir la solución $x = -2a + 4$, $y = a$. Siendo así, $T^{-1}(8, 12) = \{(-2a + 4, a): a \in \mathbf{R}\}$.

c) Hacemos $T(x, y) = (1, 2)$ llegando al sistema

$$2x + 4y = 1 \qquad 3x + 6y = 2$$

El sistema no tiene solución, luego $T^{-1}(1, 2) = \emptyset$, el conjunto vacío.

d) No, ya que, por ejemplo, $(1, 2)$ no tiene preimagen.

9.43. Sean $B = \{v_1, v_2, v_3\}$ una base de V y $B' = \{u_1, u_2\}$ una de U. Sea $T: V \to U$ lineal. Supóngase, además,

$$\begin{aligned} T(v_1) &= a_1 u_1 + a_2 u_2 \\ T(v_2) &= b_1 u_1 + b_2 u_2 \\ T(v_3) &= c_1 u_1 + c_2 u_2 \end{aligned} \qquad \text{y} \qquad A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

Mostrar que para cualquier $v \in V$, $A[v]_B = [T(v)]_{B'}$ (donde los vectores en K^2 y K^3 son vectores columna).

Supongamos $v = k_1 v_1 + k_2 v_2 + k_3 v_3$; entonces $[v]_B = [k_1, k_2, k_3]^T$. Asimismo,

$$\begin{aligned} T(v) &= k_1 T(v_1) + k_2 T(v_2) + k_3 T(v_3) = \\ &= k_1(a_1 u_1 + a_2 u_2) + k_2(b_1 u_1 + b_2 u_2) + k_3(c_1 u_1 + c_2 u_2) = \\ &= (a_1 k_1 + b_1 k_2 + c_1 k_3)u_1 + (a_2 k_1 + b_2 k_2 + c_2 k_3)u_2 \end{aligned}$$

En consecuencia,

$$[T(v)]_{B'} = \begin{pmatrix} a_1 k_1 + b_1 k_2 + c_1 k_3 \\ a_2 k_1 + b_2 k_2 + c_2 k_3 \end{pmatrix}$$

Efectuando los cálculos obtenemos

$$A[v]_B = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1 k_1 + b_1 k_2 + c_1 k_3 \\ a_2 k_1 + b_2 k_2 + c_2 k_3 \end{pmatrix} = [T(v)]_{B'}$$

9.44. Sea k un escalar no nulo. Probar que una aplicación lineal T es singular si y sólo si kT es singular. Por tanto, T es singular si y sólo si lo es $-T$.

Supongamos que T es singular. En ese caso, $T(v) = 0$ para algún vector $v \neq 0$. De aquí

$$(kT)(v) = kT(v) = k0 = 0$$

y kT es, pues, singular.

Supongamos ahora que kT es singular. Entonces $(kT)(w) = 0$ para algún vector $w \neq 0$, luego

$$T(kw) = kT(w) = (kT)(w) = 0$$

Pero $k \neq 0$ y $w \neq 0$ implica $kw \neq 0$, de manera que T es también singular.

9.45. Sea E un operador lineal en V para el cual $E^2 = E$. (Se asigna el término *proyección* a tal operador.) Sean U la imagen de E y W su núcleo. Probar que: *a*) si $u \in U$, necesariamente $E(u) = u$, es decir, E es la aplicación identidad en U; *b*) si $E \neq I$, E debe ser singular, esto es, $E(v) = 0$ para algún $v \neq 0$; *c*) $V = U \oplus W$.

a) Si $u \in U$, la imagen de E, entonces $E(v) = u$ para algún $v \in V$. Por consiguiente, utilizando $E^2 = E$ tenemos

$$u = E(v) = E^2(v) = E(E(v)) = E(u)$$

b) Si $E \neq I$, para algún $v \in V$, $E(v) = u$ con $v \neq u$. Por *a*), $E(u) = u$. De esta manera,

$$E(v - u) = E(v) - E(u) = u - u = 0 \quad \text{donde} \quad v - u \neq 0$$

c) Probemos primero que $V = U + W$. Sea $v \in V$. Hacemos $u = E(v)$ y $w = v - E(v)$. En tal caso,

$$v = E(v) + v - E(v) = u + w$$

Por definición, $u = E(v) \in U$, la imagen de E. Mostremos ahora que $w \in W$, el núcleo de E:

$$E(w) = E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

y así $w \in W$. De aquí $V = U + W$.

A continuación demostramos que $U \cap W = \{0\}$. Sea $v \in U \cap W$. Puesto que $v \in U$, $E(v) = v$ por *a*). Como $v \in W$, $E(v) = 0$. De este modo, $v = E(v) = 0$, luego $U \cap W = \{0\}$.

Las dos propiedades precedentes implican que $V = U \oplus W$.

9.46. Hallar la dimensión d de *a*) Hom $(\mathbf{R}^3, \mathbf{R}^2)$, *b*) Hom $(\mathbf{C}^3, \mathbf{R}^2)$, *c*) Hom $(V, \mathbf{R}^2)$, donde $V = \mathbf{C}^3$ visto como espacio vectorial sobre $\mathbf{R}$, *d*) $A(\mathbf{R}^3)$, *e*) $A(\mathbf{C}^3)$, *f*) $A(V)$, donde $V = \mathbf{C}^3$ visto como espacio vectorial sobre $\mathbf{R}$.

a) Siendo dim $\mathbf{R}^3 = 3$ y dim $\mathbf{R}^2 = 2$ tenemos (Teorema 9.11) $d = 3 \cdot 2 = 6$.

b) $\mathbf{C}^3$ es un espacio vectorial sobre $\mathbf{C}$ y $\mathbf{R}^2$, uno sobre $\mathbf{R}$; por tanto, no existe Hom $(\mathbf{C}^3, \mathbf{R}^2)$.

c) Como espacio vectorial sobre $\mathbf{R}$, $V = \mathbf{C}^3$ tiene dimensión 6. Por consiguiente (Teorema 9.11), $d = 6 \cdot 2 = 12$.

d) $A(\mathbf{R}^3) = $ Hom $(\mathbf{R}^3, \mathbf{R}^3)$ y dim $\mathbf{R}^3 = 3$, luego $d = 3^2 = 9$.

e) $A(\mathbf{C}^3) = $ Hom $(\mathbf{C}^3, \mathbf{C}^3)$ y dim $\mathbf{C}^3 = 3$, luego $d = 3^2 = 9$.

f) Como dim $V = 6$, $d = $ dim $A(V) = 6^2 = 36$.

PROBLEMAS SUPLEMENTARIOS

APLICACIONES

9.47. Determinar el número de aplicaciones distintas de $\{a, b\}$ en $\{1, 2, 3\}$.

9.48. Considérese la aplicación g que asigna a cada nombre en el conjunto {Betty, Martin, David, Alan, Rebecca} el número de letras diferentes necesarias para deletrearlo. Hallar: *a*) el gráfico de g, *b*) la imagen de g.

9.49. La Figura 9-5 es un diagrama de las aplicaciones $f: A \to B$, $g: B \to A$, $h: C \to B$, $F: B \to C$ y $G: A \to C$. Determinar si cada una de las siguientes expresiones define una aplicación compuesta y, en caso afirmativo, encontrar su dominio y su codominio: *a*) $g \circ f$, *b*) $h \circ f$, *c*) $F \circ f$, *d*) $G \circ f$, *e*) $g \circ h$, *f*) $h \circ G \circ g$.

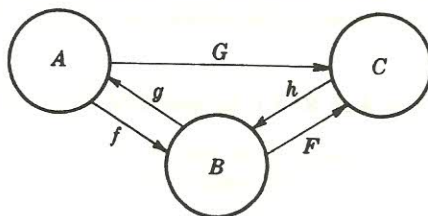


Figura 9-5.

9.50. Sean $f: \mathbf{R} \to \mathbf{R}$ y $g: \mathbf{R} \to \mathbf{R}$ las aplicaciones definidas por $f(x) = x^2 + 3x + 1$ y $g(x) = 2x - 3$. Hallar expresiones para las aplicaciones compuestas *a*) $f \circ g$, *b*) $g \circ f$, *c*) $g \circ g$, *d*) $f \circ f$.

9.51. Para cada una de las siguientes aplicaciones $f: \mathbf{R} \to \mathbf{R}$, hallar una fórmula para la aplicación inversa: *a*) $f(x) = 3x - 7$, *b*) $f(x) = x^3 + 2$.

9.52. Para cualquier aplicación $f: A \to B$, probar que $1_B \circ f = f = f \circ 1_A$.

APLICACIONES LINEALES

9.53. Comprobar que los operadores del Problema 9.40 *c*) son lineales.

a) $S(a_0 + a_1 t + \cdots + a_n t^n) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$.

b) $T(a_0 + a_1 t + \cdots + a_n t^n) = 0 + a_1 + a_2 t + \cdots + a_n t^{n-1}$.

9.54. Sean V el espacio vectorial de las matrices $n \times n$ sobre K y M una matriz arbitraria en V. Probar que las dos primeras aplicaciones $T: V \to V$ son lineales, pero la tercera no lo es (a menos que $M = 0$): *a*) $T(A) = MA$, *b*) $T(A) = MA - AM$, *c*) $T(A) = M + A$.

9.55. Hallar $T(a, b)$, donde $T: \mathbf{R}^2 \to \mathbf{R}^3$ se define según $T(1, 2) = (3, -1, 5)$ y $T(0, 1) = (2, 1, -1)$.

9.56. Dar un ejemplo de aplicación no lineal $F: V \to U$ tal que $F^{-1}(0) = \{0\}$ sin ser F inyectiva.

9.57. Mostrar que si $F: V \to U$ es lineal y aplica conjuntos independientes en conjuntos independientes, F es no singular.

9.58. Encontrar una matriz 2×2 A que aplique u_1 y u_2 en v_1 y v_2, respectivamente, siendo: *a*) $u_1 = (1, 3)^T$, $u_2 = (1, 4)^T$ y $v_1 = (-2, 5)^T$, $v_2 = (3, -1)^T$; *b*) $u_1 = (2, -4)^T$, $u_2 = (-1, 2)^T$ y $v_1 = (1, 1)^T$, $v_2 = (1, 3)^T$.

9.59. Encontrar una matriz singular 2×2 B que aplique $(1, 1)^T$ en $(1, 3)^T$.

9.60. Hallar una matriz 2×2 C que tenga un valor propio $\lambda = 3$ y aplique $(1, 1)^T$ en $(1, 3)^T$.

9.61. Sea $T: \mathbf{C} \to \mathbf{C}$ la aplicación conjugación en el cuerpo complejo $\mathbf{C}$. Esto es, $T(z) = \bar{z}$, donde $z \in \mathbf{C}$, o $T(a + bi) = a - bi$, donde $a, b \in \mathbf{R}$. *a*) Probar que T no es lineal si $\mathbf{C}$ se ve como espacio vectorial sobre sí mismo. *b*) Probar que T es lineal si $\mathbf{C}$ se ve como espacio vectorial sobre el cuerpo real $\mathbf{R}$.

9.62. Defínase $F: \mathbf{R}^2 \to \mathbf{R}^2$, según $F(x, y) = (3x + 5y, 2x + 3y)$, y sea S el círculo unidad en $\mathbf{R}^2$. (S consta de todos los puntos que satisfacen $x^2 + y^2 = 1$.) Hallar: *a*) la imagen $F(S)$, *b*) la preimagen $F^{-1}(S)$.

9.63. Considérense la aplicación lineal $G: \mathbf{R}^3 \to \mathbf{R}^3$ definida por $G(x, y, z) = (x + y + z, y - 2z, y - 3z)$ y la esfera unidad S_2 en $\mathbf{R}^3$ que consiste en los puntos que satisfacen $x^2 + y^2 + z^2 = 1$. Hallar: *a*) $G(S_2)$, *b*) $G^{-1}(S_2)$.

9.64. Sean H el plano $x + 2y - 3z = 4$ en $\mathbf{R}^3$ y G la aplicación lineal del Problema 9.63. Hallar: *a*) $G(H)$, *b*) $G^{-1}(H)$.

NUCLEO E IMAGEN DE APLICACIONES LINEALES

9.65. Para la siguiente aplicación lineal G, encontrar una base y la dimensión de i) la imagen de G, ii) el núcleo de G: $G: \mathbf{R}^3 \to \mathbf{R}^2$ definida por $G(x, y, z) = (x + y, y + z)$.

9.66. Hallar una aplicación lineal $F: \mathbf{R}^3 \to \mathbf{R}^3$ cuya imagen esté generada por (1, 2, 3) y (4, 5, 6).

9.67. Hallar una aplicación lineal $F: \mathbf{R}^4 \to \mathbf{R}^3$ cuyo núcleo esté generado por (1, 2, 3, 4) y (0, 1, 1, 1).

9.68. Sea $F: V \to U$ lineal. Mostrar que: *a*) la imagen de cualquier subespacio de V es un subespacio de U, *b*) la preimagen de cualquier subespacio de U es un subespacio de V.

9.69. Cada una de las matrices escritas a continuación determina una aplicación lineal de $\mathbf{R}^4$ en $\mathbf{R}^3$:

$$a)\quad A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{pmatrix} \qquad b)\quad B = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{pmatrix}$$

Encontrar una base y la dimensión de la imagen U y el núcleo W de dichas aplicaciones.

9.70. Considérense el espacio vectorial V de los polinomios reales $f(t)$ de grado 10 o menor y la aplicación lineal $\mathbf{D}^4: V \to V$ definida por $d^4 f/dt^4$, o sea, la cuarta derivada. Hallar una base y la dimensión de *a*) la imagen de $\mathbf{D}^4$, *b*) el núcleo de $\mathbf{D}^4$.

OPERACIONES CON APLICACIONES LINEALES

9.71. Difínanse $F: \mathbf{R}^3 \to \mathbf{R}^2$ y $G: \mathbf{R}^3 \to \mathbf{R}^2$ según $F(x, y, z) = (y, x + z)$ y $G(x, y, z) = (2z, x - y)$. Encontrar expresiones que definan las aplicaciones $F + G$ y $3F - 2G$.

9.72. Defínase $H: \mathbf{R}^2 \to \mathbf{R}^2$ por $H(x, y) = (y, 2x)$. Empleando las aplicaciones F y G del Problema 9.71, hallar expresiones que definan: *a*) $H \circ F$ y $H \circ G$, *b*) $F \circ H$ y $G \circ H$, *c*) $H \circ (F + G)$ y $H \circ F + H \circ G$.

9.73. Mostrar que las aplicaciones F, G y H que siguen son linealmente independientes:

a) $F, G, H \in \text{Hom}(\mathbf{R}^2, \mathbf{R}^2)$ definidas por $F(x, y) = (x, 2y)$, $G(x, y) = (y, x + y)$, $H(x, y) = (0, x)$.

b) $F, G, H \in \text{Hom}(\mathbf{R}^3, \mathbf{R})$ definidas por $F(x, y, z) = x + y + z$, $G(x, y, z) = y + z$, $H(x, y, z) = x - z$.

9.74. Para $F, G \in \text{Hom}(V, U)$, probar que rango $(F + G) \leq$ rango F + rango G. (Aquí V tiene dimensión finita.)

9.75. Sean $F: V \to U$ y $G: U \to V$ lineales. Probar que si F y G son no singulares, $G \circ F$ es no singular. Dar un ejemplo en el que $G \circ F$ sea no singular siendo G singular.

9.76. Demostrar que Hom (V, U) satisface todos los axiomas que definen un espacio vectorial. Es decir, demostrar el Teorema 9.10.

ALGEBRA DE OPERADORES LINEALES

9.77. Supóngase que S y T son operadores lineales en V y que S es no singular. Supóngase, asimismo, que V tiene dimensión finita. Mostrar que rango (ST) = rango (TS) = rango T.

9.78. Supóngase $V = U \oplus W$. Sean E_1 y E_2 operadores lineales en V definidos por $E_1(v) = u$, $E_2(v) = w$, donde $v = u + w$, $u \in U$, $w \in W$. Probar que: *a*) $E_1^2 = E_1$ y $E_2^2 = E_2$, o sea, que E_1 y E_2 son «proyecciones»; *b*) $E_1 + E_2 = I$, la aplicación identidad; *c*) $E_1 E_2 = \mathbf{0}$ y $E_2 E_1 = \mathbf{0}$.

9.79. Sean E_1 y E_2 operadores lineales en V que satisfacen las condiciones *a*), *b*) y *c*) del Problema 9.78. Demostrar que $E_2: V = \text{Im } E_1 \oplus \text{Im } E_2$.

9.80. Mostrar que si los operadores lineales S y T son invertibles, necesariamente lo es ST y $(ST)^{-1} = T^{-1}S^{-1}$.

9.81. Sean V de dimensión finita y T un operador lineal en V tal que rango T^2 = rango T. Probar que Ker $T \cap \text{Im } T = \{0\}$.

9.82. ¿Cuáles entre los siguientes enteros pueden ser dimensión de un álgebra $A(V)$ de aplicaciones lineales: 5, 9, 18, 25, 31, 36, 44, 64, 88, 100?

9.83. Se dice que un álgebra A tiene un elemento identidad 1 si $1 \cdot a = a \cdot 1 = a$ para todo $a \in A$. Mostrar que $A(V)$ tiene un elemento identidad.

9.84. Hallar la dimensión de $A(V)$, donde: *a*) $V = \mathbf{R}^4$, *b*) $V = \mathbf{C}^4$, *c*) $V = \mathbf{C}^4$ visto como espacio vectorial sobre $\mathbf{R}$, *d*) V = polinomios de grado ≤ 10.

PROBLEMAS VARIOS

9.85. Supóngase que $T: K^n \to K^m$ es una aplicación lineal. Sean $\{e_1, \ldots, e_n\}$ la base usual de K^n y A la matriz $m \times n$ cuyas columnas son los vectores $T(e_1), \ldots, T(e_n)$, respectivamente. Probar que, para todo vector $v \in K^n$, $T(v) = Av$, donde v se escribe como un vector columna.

9.86. Supóngase que $F: V \to U$ es lineal y que k es un escalar no nulo. Probar que las aplicaciones F y kF tienen el mismo núcleo y la misma imagen.

9.87. Demostrar que si $F: V \to U$ es suprayectiva, necesariamente dim $U \leq$ dim V. Determinar todas las aplicaciones $T: \mathbf{R}^3 \to \mathbf{R}^4$ que sean suprayectivas.

9.88. Sean $T: V \to U$ lineal y W un subespacio de V. La restricción de T a W es la aplicación $T_W: W \to U$ definida por $T_W(w) = T(w)$ para todo $w \in W$. Demostrar las siguientes afirmaciones: *a*) T_W es lineal. *b*) Ker T_W = Ker $T \cap W$. *c*) Im $T_W = T(w)$.

9.89. Se dice que dos operadores $S, T \in A(V)$ son *similares* si existe un operador invertible $P \in A(V)$ para el que $S = P^{-1}TP$. Demostrar las siguientes afirmaciones: *a*) La similaridad de operadores es una relación de equivalencia. *b*) Los operadores similares tienen el mismo rango (cuando V es de dimensión finita).

9.90. Sean v y w elementos de un espacio vectorial real V. El *segmento rectilíneo* L desde v hasta $v + w$ se define como el conjunto de vectores $v + tw$ para $0 \leq t \leq 1$. (Véase la Figura 9-6.)

a) Mostrar que el segmento rectilíneo L entre dos vectores v y u consiste en los puntos:
i) $(1 - t)v + tu$ para $0 \leq t \leq 1$, ii) $t_1 v + t_2 u$ para $t_1 + t_2 = 1$, $t_1 \geq 0$, $t_2 \geq 0$.

b) Sea $F: V \to U$ lineal. Mostrar que la imagen $F(L)$ de un segmento rectilíneo L en V es un segmento rectilíneo en U.

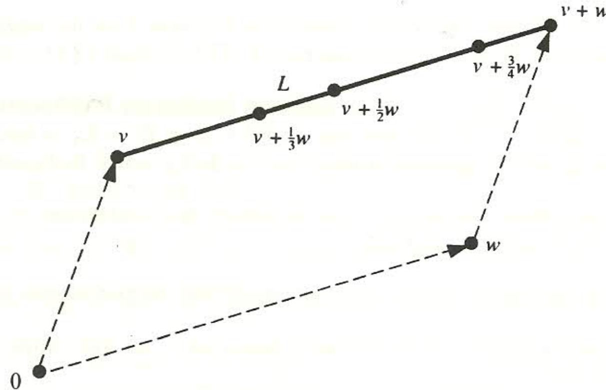


Figura 9-6.

9.91. Se dice que un subconjunto X de un espacio vectorial V es *convexo* si el segmento rectilíneo L entre dos puntos (vectores) cualesquiera $P, Q \in X$ está contenido en X.

a) Probar que la intersección de conjuntos convexos es convexa.

b) Supóngase que $F: V \to U$ es lineal y X convexo. Probar que $F(X)$ es convexo.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

9.47. Nueve.

9.48. *a*) {(Betty, 4), (Martin, 6), (David, 4), (Alan, 3), (Rebecca, 5)}. *b*) Imagen de $g = \{3, 4, 5, 6\}$.

9.49. *a*) $(g \circ f): A \to A$, *b*) no, *c*) $(F \circ f): A \to C$, *d*) no, *e*) $(g \circ h): C \to A$,
f) $(h \circ G \circ g): B \to B$.

9.50. *a*) $(f \circ g)(x) = 4x^2 - 6x + 1$ *c*) $(g \circ g)(x) = 4x - 9$
b) $(g \circ f)(x) = 2x^2 + 6x - 1$ *d*) $(f \circ f)(x) = x^4 + 6x^3 + 14x^2 + 15x + 5$

9.51. *a*) $f^{-1}(x) = (x + 7)/3$, *b*) $f^{-1}(x) = \sqrt[3]{x - 2}$.

9.55. $T(a, b) = (-a + 2b, -3a + b, 7a - b)$.

9.56. Tomar $V = U = \mathbf{R}^2$ y $F(x, y) = (x^2, y^2)$.

9.58. *a*) $\begin{pmatrix} -17 & 5 \\ 23 & -6 \end{pmatrix}$. *b*) No existe, pues u_1, u_2 son dependientes y v_1, v_2 no lo son.

9.59. $\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}$. [*Indicación:* Envíese $(0, 1)^T$ a $(0, 0)^T$.]

9.60. $\begin{pmatrix} -2 & 3 \\ 0 & 3 \end{pmatrix}$. [*Indicación:* Envíese $(0, 1)^T$ a $(0, 3)^T$.]

9.62. *a*) $13x^2 - 42xy + 34y^2 = 1$, *b*) $13x^2 + 42xy + 24y^2 = 1$.

9.63. *a*) $x^2 - 8xy + 26y^2 + 6xz - 38yz + 14z^2 = 1$, *b*) $x^2 + 2xy + 3y^2 + 2xz - 8yz + 14z^2 = 1$.

9.64. *a*) $x - y + 2z = 4$, *b*) $x - 12z = 4$.

9.65. i) $(1, 0)$, $(0, 1)$, rango $G = 2$; ii) $(1, -1, 1)$, nulidad $G = 1$.

9.66. $F(x, y, z) = (x + 4y, 2x + 5y, 3x + 6y)$.

9.67. $F(x, y, z, t) = (x + y - z, 3x + y - t, 0)$.

9.69. *a*) $\{(1, 2, 1), (0, 1, 1)\}$ es una base de Im A; dim (Im A) = 2.
$\{(4, -2, -5, 0), (1, -3, 0, 5)\}$ es una base de Ker A; dim (Ker A) = 2.
b) Im $B = \mathbf{R}^3$; $\{(-1, \frac{2}{3}, 1, 1)\}$ es una base de Ker B; dim (Ker B) = 1.

9.70. *a*) $1, t, \ldots, t^6$; rango $\mathbf{D}^4 = 7$; *b*) $1, t, t^2, t^3$; nulidad $\mathbf{D}^4 = 4$.

9.71. $(F + G)(x, y, z) = (y + 2z, 2x - y + z)$, $(3F - 2G)(x, y, z) = (3y - 4z, x + 2y + 3z)$.

9.72. *a*) $(H \circ F)(x, y, z) = (x + y, 2y)$, $(H \circ G)(x, y, z) = (x - y, 4z)$. *b*) No están definidas.
c) $(H \circ (F + G))(x, y, z) = (H \circ F + H \circ G)(x, y, z) = (2x - y + z, 2y + 4z)$.

9.82. Los cuadrados: 9, 25, 36, 64, 100.

9.84. *a*) 16, *b*) 16, *c*) 64, *d*) 121.

CAPITULO 10

Matrices y aplicaciones lineales

10.1. INTRODUCCION

Supongamos que $S = \{u_1, u_2, \ldots, u_n\}$ es una base de un espacio vectorial V sobre un cuerpo K, y para $v \in V$ supongamos que

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

Entonces el vector coordenado de v relativo a la base S, que escribiremos como vector columna a menos que se especifique o sobrentienda lo contrario, se denota y define por

$$[v]_S = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix} = [a_1, a_2, \ldots, a_n]^T$$

Recuérdese (Ejemplo 9.4) que la aplicación $v \to [v]_S$, determinada por la base S, es un isomorfismo entre V y el espacio K^n.

En este capítulo mostraremos que también existe un isomorfismo, determinado por la base S, entre el álgebra $A(V)$ de operadores lineales en V y el álgebra $\mathbf{M}$ de matrices n-cuadradas sobre K. Así todo operador lineal $T: V \to V$ corresponderá a una matriz n-cuadrada $[T]_S$ determinada por la base S.

En el capítulo se intentará, asimismo, contestar a la pregunta de si un operador lineal T puede o no representarse por una matriz diagonal.

10.2. REPRESENTACION MATRICIAL DE UN OPERADOR LINEAL

Sea T un operador lineal en un espacio vectorial V sobre un cuerpo K y supongamos que $S = \{u_1, u_2, \ldots, u_n\}$ es una base de V. Ahora $T(u_1), \ldots, T(u_n)$ son vectores en V y por tanto cada uno de ellos es combinación lineal de los vectores de la base S, digamos

$$\begin{aligned} T(u_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ T(u_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(u_n) &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n \end{aligned}$$

Es pertinente la siguiente definición.

Definición: La traspuesta de la matriz de los coeficientes anterior, denotada por $m_S(T)$ o $[T]_S$, se denomina la *representación matricial* de T relativa a la base S o, simplemente, la matriz de T en la base S; esto es,

$$[T]_S = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

(El subíndice S puede omitirse si la base S viene dada implícitamente.)

Nota: Utilizando la notación de vectores (columna) coordenados, la representación matricial de T puede expresarse también

$$m(T) = [T] = ([T(u_1)], [T(u_2)], \ldots, [T(u_n)])$$

es decir, las columnas de $m(T)$ son los vectores coordenados $[T(u_1)], \ldots, [T(u_n)]$.

EJEMPLO 10.1

a) Sea V el espacio vectorial de los polinomios en t sobre $\mathbf{R}$ de grado $\leqslant 3$ y $\mathbf{D}: V \to V$ el operador de derivación definido por $\mathbf{D}(p(t)) = d(p(t))/dt$. Calculemos la matriz de $\mathbf{D}$ en la base $\{1, t, t^2, t^3\}$.

$$\begin{aligned} \mathbf{D}(1) &= 0 &&= 0 - 0t + 0t^2 + 0t^3 \\ \mathbf{D}(t) &= 1 &&= 1 + 0t + 0t^2 + 0t^3 \\ \mathbf{D}(t^2) &= 2t &&= 0 + 2t + 0t^2 + 0t^3 \\ \mathbf{D}(t^3) &= 3t^2 &&= 0 + 0t + 3t^2 + 0t^3 \end{aligned} \qquad \text{y} \qquad [\mathbf{D}] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

[Nótese que los vectores coordenados de $\mathbf{D}(1)$, $\mathbf{D}(t)$, $\mathbf{D}(t^2)$ y $\mathbf{D}(t^3)$ son las columnas, no las filas, de $[\mathbf{D}]$.]

b) Consideremos el operador lineal $F: \mathbf{R}^2 \to \mathbf{R}^2$ definido por $F(x, y) = (4x - 2y, 2x + y)$ y las bases de $\mathbf{R}^2$:

$$S = \{u_1 = (1, 1), u_2 = (-1, 0)\} \qquad \text{y} \qquad E = \{e_1 = (1, 0), e_2 = (0, 1)\}$$

Tenemos

$$F(u_1) = F(1, 1) = (2, 3) = 3(1, 1) + (-1, 0) = 3u_1 + u_2$$
$$F(u_2) = F(-1, 0) = (-4, -2) = -2(1, 1) + 2(-1, 0) = -2u_1 + 2u_2$$

Por consiguiente, $[F]_S = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$ es la representación matricial de F relativa a la base S. Tenemos también

$$F(e_1) = F(1, 0) = (4, 2) = 4e_1 + 2e_2$$
$$F(e_2) = F(0, 1) = (-2, 1) = -2e_1 + e_2$$

En consecuencia, $[F]_E = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}$ es la representación matricial de F relativa a la base usual E.

c) Consideremos cualquier matriz n-cuadrada A sobre K (que define una aplicación lineal $A: K^n \to K^n$) y la base usual $E = \{e_i\}$ de K^n. En ese caso, $Ae_1, Ae_2, \ldots, Ae_n$ son precisamente las columnas de A (Ejemplo 9.5) y sus coordenadas relativas a la base usual E son los propios vectores. De acuerdo con ello,

$$[A]_E = A$$

o sea, respecto a la base usual E, la representación matricial de una aplicación matricial A es la misma matriz A.

El algoritmo que ahora exponemos se usará para calcular representaciones matriciales.

Algoritmo 10.2

Dados un operador lineal T en V y una base $S = \{u_1, \ldots, u_n\}$ de V, este algoritmo halla la matriz $[T]_S$ que representa T en la base S.

Paso 1. Repetir para cada vector u_k de la base S:

a) Hallar $T(u_k)$.

b) Escribir $T(u_k)$ como combinación lineal de los vectores $u_1, \ldots, u_n$ para obtener las coordenadas de $T(u_k)$ en la base S.

Paso 2. Construir la matriz $[T]_S$ cuyas columnas son los vectores coordenados $[T(u_k)]_S$ obtenidos en el Paso 1.

Paso 3. Salir.

Nota: Obsérvese que el Paso 1 *b*) se repite para cada vector u_k de la base. En consecuencia, puede ser útil efectuar previamente:

Paso 0. Encontrar una fórmula para las coordenadas de un vector arbitrario v relativas a la base S.

Nuestro primer teorema, demostrado en el Problema 10.10, nos dice que la «acción» de un operador T sobre un vector v es preservada por su representación matricial.

Teorema 10.1: Sean $S = \{u_1, u_2, \ldots, u_n\}$ una base de V y T un operador lineal en V. Entonces, para todo vector $v \in V$, $[T]_S[v]_S = [T(v)]_S$.

Esto es, si multiplicamos el vector coordenado de v por la representación matricial de T, conseguimos el vector coordenado de $T(v)$.

EJEMPLO 10.2. Consideremos el operador de derivación $\mathbf{D}: V \to V$ del Ejemplo 10.1 *a*). Sea

$$p(t) = a + bt + ct^2 + dt^3 \quad \text{de modo que} \quad \mathbf{D}(p(t)) = b + 2ct + 3dt^2$$

Por tanto, respecto a la base $\{1, t, t^2, t^3\}$,

$$[p(t)] = [a, b, c, d]^T \quad \text{y} \quad [\mathbf{D}(p(t))] = [b, 2c, 3d, 0]^T$$

Probemos que el Teorema 10.1 es válido aquí:

$$[\mathbf{D}][p(t)] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b \\ 2c \\ 3d \\ 0 \end{pmatrix} = [\mathbf{D}(p(t))]$$

Hemos asociado una matriz $[T]$ a cada T en $A(V)$, el álgebra de los operadores lineales en V. En virtud de nuestro primer teorema, la acción de un operador individual T es preservada por esta representación. Los próximos dos teoremas (demostrados en los Problemas 10.11 y 10.12, respectivamente) nos dicen que las tres operaciones básicas con estos operadores,

i) suma,

ii) producto por un escalar,

iii) composición,

también se preservan.

Teorema 10.2: Sean $S = \{u_1, u_2, \ldots, u_n\}$ una base de un espacio vectorial V sobre K y $\mathbf{M}$ el álgebra de las matrices n-cuadradas sobre K. En tal caso, la aplicación $m: A(V) \to \mathbf{M}$ definida por $m(T) = [T]_S$ es un isomorfismo entre espacios vectoriales. Es decir, para todo par de operadores F, $G \in A(V)$ y todo $k \in K$ tenemos

i) $m(F + G) = m(F) + m(G)$, o sea, $[F + G] = [F] + [G]$.

ii) $m(kF) = km(F)$, o sea, $[kF] = k[F]$.

iii) m es inyectiva y suprayectiva.

Teorema 10.3: Para todo par de operadores lineales G, $F \in Q(v)$,

$$m(G \circ F) = m(G)m(F) \quad \text{esto es,} \quad [G \circ F] = [G][F]$$

(Aqui $G \circ F$ denota la composición de las aplicaciones G y F.)

Ilustremos los teoremas precedentes en el caso dim $V = 2$. Supongamos que $\{u_1, u_2\}$ es una base de V y que F y G son operadores lineales en V para los que

$$F(u_1) = a_1 u_1 + a_2 u_2 \qquad G(u_1) = c_1 u_1 + c_2 u_2$$
$$F(u_2) = b_1 u_1 + b_2 u_2 \qquad G(u_2) = d_1 u_1 + d_2 u_2$$

Entonces $$[F] = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad \text{y} \quad [G] = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}$$

Tenemos

$$(F+G)(u_1) = F(u_1) + G(u_1) = a_1 u_1 + a_2 u_2 + c_1 u_1 + c_2 u_2 =$$
$$= (a_1 + c_1)u_1 + (a_2 + c_2)u_2$$
$$(F+G)(u_2) = F(u_2) + G(u_2) = b_1 u_1 + b_2 u_2 + d_1 u_1 + d_2 u_2 =$$
$$= (b_1 + d_1)u_1 + (b_2 + d_2)u_2$$

Siendo así, $$[F+G] = \begin{pmatrix} a_1 + c_1 & b_1 + d_1 \\ a_2 + c_2 & b_2 + d_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} + \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = [F] + [G]$$

Asimismo, para $k \in K$,

$$(kF)(u_1) = kF(u_1) = k(a_1 u_1 + a_2 u_2) = ka_1 u_1 + ka_2 u_2$$
$$(kF)(u_2) = kF(u_2) = k(b_1 u_1 + b_2 u_2) = kb_1 u_1 + kb_2 u_2$$

De este modo,

$$[kF] = \begin{pmatrix} ka_1 & kb_1 \\ ka_2 & kb_2 \end{pmatrix} = k\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = k[F]$$

Finalmente,

$$(G \circ F)(u_1) = G(F(u_1)) = G(a_1 u_1 + a_2 u_2) = a_1 G(u_1) + a_2 G(u_2) =$$
$$= a_1(c_1 u_1 + c_2 u_2) + a_2(d_1 u_1 + d_2 u_2) =$$
$$= (a_1 c_1 + a_2 d_1)u_1 + (a_1 c_2 + a_2 d_2)u_2$$

$$(G \circ F)(u_2)) = G(F(u_2) = G(b_1 u_1 + b_2 u_2) = b_1 G(u_1) + b_2 G(u_2) =$$
$$= b_1(c_1 u_1 + c_2 u_2) + b_2(d_1 u_1 + d_2 u_2) =$$
$$= (b_1 c_1 + b_2 d_1)u_1 + (b_1 c_2 + b_2 d_2)u_2$$

En consecuencia,

$$[G \circ F] = \begin{pmatrix} a_1 c_1 + a_2 d_1 & b_1 c_1 + b_2 d_1 \\ a_1 c_2 + a_2 d_2 & b_1 c_2 + b_2 d_2 \end{pmatrix} = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = [G][F]$$

10.3. CAMBIO DE BASE Y OPERADORES LINEALES

La discusión anterior muestra que podemos representar un operador lineal por una matriz una vez hayamos elegido una base. Surge de forma natural la siguiente pregunta: ¿Cómo varía

nuestra representación si seleccionamos otra base? Para responder a esta pregunta, empezamos por recordar una definición y algunos hechos.

Definición: Sean $S = \{u_1, u_2, \ldots, u_n\}$ una base de V y $S' = \{v_1, v_2, \ldots, v_n\}$ otra base. Supongamos que, para $i = 1, 2, \ldots, n$,

$$v_i = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n$$

La traspuesta P de la matriz de coeficientes precedente se llama matriz de cambio de base (o matriz de transición) desde la «antigua» base S hasta la nueva «base» S'.

Hecho 1. La matriz de cambio de base P es invertible, siendo su inversa P^{-1} la matriz de cambio de base desde la S' hasta la S.

Hecho 2. Sea P la matriz de cambio de base desde la usual E de K^n hasta otra base S. Entonces P es la matriz cuyas columnas con precisamente los elementos de S.

Hecho 3. Sea P la matriz de cambio de base desde una base S hasta otra S' en V. En tal caso (Teorema 5.27), para todo vector $v \in V$,

$$P[v]_{S'} = [v]_S \qquad \text{y} \qquad P^{-1}[v]_S = [v]_{S'}$$

(De este modo, P^{-1} transforma las coordenadas de v en la «antigua» base S en las de v en la «nueva» base S'.)

El teorema enunciado a continuación, demostrado en el Problema 10.19, contesta a la pregunta que se hizo con anterioridad, o sea, muestra cómo se ve afectada la representación matricial de un operador lineal por un cambio de base.

Teorema 10.4: Sea P la matriz de cambio de base desde una base S hasta otra S' en un espacio vectorial V. Para todo operador lineal T en V,

$$[T]_{S'} = P^{-1}[T]_S P$$

En otras palabras, si A es la matriz que representa T en la base S, $B = P^{-1}AP$ es la que lo representa en una nueva base S', donde P es la matriz de cambio de base desde S hasta S'.

EJEMPLO 10.3. Consideremos las bases de $\mathbf{R}^2$:

$$E = \{e_1 = (1, 0), e_2 = (0, 1)\} \qquad \text{y} \qquad S = \{u_1 = (1, -2), u_2 = (2, -5)\}$$

Como E es la base usual de $\mathbf{R}^2$, escribimos los vectores de la base S como columnas para obtener la matriz de cambio de base P de E a S:

$$P = \begin{pmatrix} 1 & 2 \\ -2 & -5 \end{pmatrix}$$

Consideremos el operador lineal F en $\mathbf{R}^2$ definido por $F(x, y) = (2x - 3y, 4x + y)$. Se tiene

$$\begin{aligned} F(e_1) &= F(1, 0) = (2, 4) = 2e_1 + 4e_2 \\ F(e_2) &= F(0, 1) = (-3, 1) = -3e_1 + e_2 \end{aligned} \qquad \text{y por consiguiente} \quad A = \begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}$$

es la representación matricial de F relativa a la base usual E. Según el Teorema 10.4,

$$B = P^{-1}AP = \begin{pmatrix} 5 & 2 \\ -2 & -1 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & -5 \end{pmatrix} = \begin{pmatrix} 44 & 101 \\ -18 & -41 \end{pmatrix}$$

es la representación matricial de F relativa a la base S.

Nota: Supongamos que $P = (a_{ij})$ es cualquier matriz n-cuadrada invertible sobre un cuerpo K y que $S = \{u_1, u_2, \ldots, u_n\}$ es una base de un espacio vectorial V sobre K. En tal caso, los n vectores

$$v_i = a_{1i}u_1 + a_{2i}u_2 + \cdots + a_{ni}u_n \qquad i = 1, 2, \ldots, n$$

son linealmente independientes y por ende constituyen otra base S' de V. Aún más, P es la matriz de cambio de base desde la base S hasta la S'. En consecuencia, si A es una representación matricial de un operador lineal T en V, la matriz $B = P^{-1}AP$ será también una representación matricial de T.

SIMILARIDAD Y OPERADORES LINEALES

Supongamos que A y B son matrices cuadradas para las que existe una matriz invertible P tal que $B = P^{-1}AP$. Entonces (Sección 4.13) se dice que B es *similar* a A, o que se obtiene de A mediante una *transformación de similaridad*. De acuerdo con el Teorema 10.4 y la nota anterior, disponemos del siguiente resultado básico.

Teorema 10.5: Dos matrices A y B representan el mismo operador lineal T si y sólo si son similares entre sí.

Esto es, todas las representaciones matriciales de un operador lineal T forman una clase de equivalencia de matrices similares.

Supongamos ahora que f es una función definida sobre las matrices cuadradas que asigna el mismo valor a las matrices similares; es decir, $f(A) = f(B)$ siempre que A sea similar a B. Entonces f induce una función, que también denotaremos por f, definida sobre los operadores lineales T de la forma natural: $f(T) = f([T]_S)$, donde S es cualquier base. La función está bien definida, según el Teorema 10.5. Tres importantes ejemplos de tales funciones son:

1. El determinante. 2. La traza. 3. El polinomio característico.

El determinante, la traza y el polinomio característico de un operador lineal T están, pues, bien definidos:

EJEMPLO 10.4. Sea F el operador lineal en $\mathbf{R}^2$ definido por $F(x, y) = (2x - 3y, 4x + y)$. Por el Ejemplo 10.3, la representación matricial de T relativa a la base usual de $\mathbf{R}^2$ es

$$A = \begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}$$

De acuerdo con esto:

i) $\det(T) = \det(A) = 2 + 12 = 14$ es el determinante de T.

ii) $\operatorname{tr} T = \operatorname{tr} A = 2 + 1 = 3$ es la traza de T.

iii) $\Delta_T(t) = \Delta_A(t) = t^2 - 3t + 14$ es el polinomio característico de T.

También según el Ejemplo 10.3, otra representación matricial de T es la matriz

$$B = \begin{pmatrix} 44 & 101 \\ -18 & -41 \end{pmatrix}$$

Usando ésta obtenemos:

i) $\det(T) = \det(A) = -1804 + 1818 = 14$ es el determinante de T.

ii) $\operatorname{tr} T = \operatorname{tr} A = 44 - 41 = 3$ es la traza de T.

iii) $\Delta_T(t) = \Delta_B(t) = t^2 - 3t + 14$ es el polinomio característico de T.

Como cabía esperar, ambas matrices proporcionan los mismos resultados.

10.4. DIAGONALIZACION DE OPERADORES LINEALES

Un operador lineal T en un espacio vectorial V se dice *diagonalizable* si puede representarse por una matriz diagonal D. Así pues, T es diagonalizable si y sólo si existe una base $S = \{u_1, u_2, \ldots, u_n\}$ de V para la cual

$$\begin{aligned} T(u_1) &= k_1 u_1 \\ T(u_2) &= \quad k_2 u_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ T(u_n) &= \qquad\quad k_n u_n \end{aligned}$$

En tal caso, T se representa por la matriz diagonal

$$D = \operatorname{diag}(k_1, k_2, \ldots, k_n)$$

respecto a la base S.

La observación precedente nos conduce a las definiciones y teoremas que enseguida escribimos, análoga a las definiciones y teoremas para matrices que se discutieron en el Capítulo 8.

Un escalar $\lambda \in K$ se denomina *valor propio* de T si existe un vector no nulo $v \in V$ para el cual

$$T(v) = \lambda v$$

Todo vector que satisfaga esta relación recibe el nombre de *vector propio* de T *perteneciente* al valor propio λ. El conjunto E_λ de todos los vectores semejantes es un subespacio de V, llamado el *espacio propio* de λ. (Alternativamente, λ es un valor propio de T si $\lambda I - T$ es singular y, cuando así sea, E_λ es el núcleo de $\lambda I - T$.)

Disponemos de los siguientes teoremas.

Teorema 10.6: T puede representarse por una matriz diagonal D (o T es diagonalizable) si y sólo si existe una base S de V que consiste en vectores propios de T, en cuyo caso los elementos diagonales de D son los valores propios correspondientes.

Teorema 10.7: Vectores propios no nulos $u_1, u_2, \ldots, u_r$ de T, pertenecientes, respectivamente, a valores propios distintos $\lambda_1, \lambda_2, \ldots, \lambda_r$, son linealmente independientes. (Véase la demostración en el Problema 10.26.)

Teorema 10.8: T es una raíz de su polinomio característico $\Delta(t)$.

Teorema 10.9: El escalar λ es un valor propio de T si y sólo si es una raíz del polinomio característico $\Delta(t)$ de T.

Teorema 10.10: La multiplicidad geométrica de un valor propio λ de T no excede su multiplicidad algebraica. (Véase la demostración en el Problema 10.27.)

Teorema 10.11: Supongamos que A es una representación matricial de T. Entonces T es diagonalizable si y sólo si lo es A.

Nota: El Teorema 10.11 reduce el estudio de la diagonalización de un operador lineal T al de la diagonalización de una matriz A, que se discutió en detalle en el Capítulo 8.

EJEMPLO 10.5

a) Sean V el espacio vectorial de las funciones reales, para el que $S = \{\text{sen } \theta, \cos \theta\}$ es una base, y $\mathbf{D}$ el operador de derivación en V. En tal caso,

$$\mathbf{D}(\text{sen } \theta) = \cos\theta = 0(\text{sen } \theta) + 1(\cos\theta)$$
$$\mathbf{D}(\cos\theta) = -\text{sen } \theta = -1(\text{sen } \theta) + 0(\cos\theta)$$

De aquí que $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ sea la representación matricial de $\mathbf{D}$ en la base S. Por tanto,

$$\Delta(t) = t^2 - (\text{tr } A)t + |A| = t^2 + 1$$

es el polinomio característico tanto de A como de $\mathbf{D}$. De esta manera, A y $\mathbf{D}$ no tienen valores propios (reales) y en particular $\mathbf{D}$ no es diagonalizable.

b) Consideremos las funciones $e^{a_1 t}, e^{a_2 t}, \ldots, e^{a_r t}$, donde $a_1, a_2, \ldots, a_r$ son números reales diferentes. Sea $\mathbf{D}$ el operador de derivación; por consiguiente, $\mathbf{D}(e^{a_k t}) = a_k e^{a_k t}$. En consecuencia, las funciones $e^{a_k t}$ son vectores propios de $\mathbf{D}$ pertenecientes a valores propios distintos. Las funciones son, pues, linealmente independientes, de acuerdo con el Teorema 10.7.

c) Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ el operador lineal que gira cada vector $v \in \mathbf{R}^2$ un ángulo $\theta = 90°$ (como se muestra en la Figura 10-1). Nótese que ningún vector no nulo es transformado en un múltiplo de sí mismo, por lo que T no tiene valores propios reales ni por ende vectores propios.

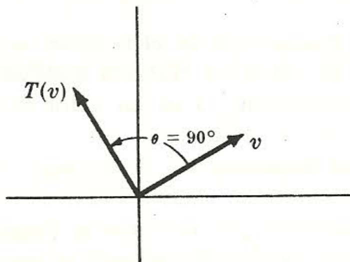


Figura 10-1.

Ahora el *polinomio mínimo* $m(t)$ de un operador lineal T se define, con independencia de la teoría de matrices, como el polinomio normalizado de menor grado que tenga T como cero. No obstante, para todo polinomio $f(t)$,

$$f(T) = \mathbf{0} \quad \text{si y sólo si} \quad f(A) = 0$$

donde A es cualquier representación matricial de T. En consecuencia, T y A tienen el mismo polinomio mínimo, por lo que todos los teoremas del Capítulo 8 referentes al polinomio mínimo de una matriz rigen también para el polinomio mínimo de un operador lineal T.

10.5. MATRICES Y APLICACIONES LINEALES GENERALES

Consideremos, por último, el caso general de las aplicaciones lineales de un espacio vectorial en otro. Sean V y U espacios vectoriales sobre el mismo cuerpo K y, digamos, $\dim V = m$ y $\dim U = n$. Sean además $\mathbf{e} = \{v_1, v_2, \ldots, v_m\}$ y $\mathbf{f} = \{u_1, u_2, \ldots, u_n\}$ bases arbitrarias pero fijas de V y U, respectivamente.

Supongamos que $F: V \to U$ es una aplicación lineal. Entonces los vectores $F(v_1), \ldots, F(v_m)$ pertenecen a U, luego cada uno de ellos es una combinación lineal de los u_k; por ejemplo,

$$\begin{aligned} F(v_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ F(v_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F(v_m) &= a_{m1}u_n + a_{m2}u_2 + \cdots + a_{mn}u_n \end{aligned}$$

La traspuesta de la matriz de coeficientes anterior, denotada por $[F]_e^f$, se llama la *representación matricial* de F relativa a las bases $\mathbf{e}$ y $\mathbf{f}$:

$$[F]_e^f = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}$$

(Emplearemos la notación $[F]$, más sencilla, cuando las bases vengan dadas implícitamente.)

Son aplicables los siguientes teoremas.

Teorema 10.12: Para todo vector $v \in V$, $[F]_e^f [v]_e = [F(v)]_f$.

O sea, multiplicando el vector coordenado de v en la base $\mathbf{e}$ por la matriz $[F]_e^f$ obtenemos el vector coordenado de $F(v)$ en la base $\mathbf{f}$.

Teorema 10.13: La aplicación $F \mapsto [F]$ es un isomorfismo de Hom (V, U) sobre el espacio vectorial de las matrices $m \times n$ sobre K. Esto es, la aplicación es inyectiva y suprayectiva y para todo par $F, G \in \text{Hom}(V, U)$ y todo $k \in K$,

$$[F + G] = [F] + [G] \qquad \text{y} \qquad [kF] = k[F]$$

Nota: Recuérdese que toda matriz $n \times m$ A sobre K puede identificarse con la aplicación lineal de K^m en K^n dada por $v \mapsto Av$. Supongamos ahora que V y U son espacios vectoriales sobre K de dimensiones m y n, respectivamente, y que **e** es una base de V y **f** una de U. En vista del teorema precedente, identificaremos también A con la aplicación lineal $F: V \to U$ dada por $[F(v)]_f = A[v]_e$. Cabe comentar que si se dan otras bases de V y U, A se identificará con otra aplicación lineal de V en U.

Teorema 10.14: Sean **e**, **f** y **g** bases de V, U y W, respectivamente. Sean, asimismo, $F: V \to U$ y $G: U \to W$ aplicaciones lineales. Entonces

$$[G \circ F]_e^g = [G]_f^g [F]_e^f$$

Es decir, respecto a las bases apropiadas, la representación matricial de la composición de dos aplicaciones lineales es igual al producto de las representaciones matriciales de las aplicaciones individuales.

El próximo teorema muestra cómo se ve afectada la representación matricial de una aplicación lineal $F: V \to U$ por la selección de nuevas bases.

Teorema 10.15: Sea P la matriz de cambio de base desde una base **e** hasta otra **e**′ en V y sea Q la matriz de cambio de base desde una base **f** hasta otra **f**′ en U. Para toda aplicación lineal $F: V \to U$,

$$[F]_{e'}^{f'} = Q^{-1}[F]_e^f P$$

Dicho de otro modo, si A representa la aplicación lineal F respecto a las bases **e** y **f**

$$B = Q^{-1}AP$$

representa F en las nuevas bases **e**′ y **f**′.

Nuestro último teorema, demostrado en el Problema 10.34, muestra que toda aplicación lineal de un espacio en otro puede representarse por una matriz muy simple.

Teorema 10.16: Sea $F: V \to U$ lineal y, digamos, rango $F = r$. Existen bases de V y U en las que la representación matricial de F adopta la forma

$$A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

donde I es la matriz identidad r-cuadrada.

La matriz A anterior se conoce como la forma *normal* o *canónica* de la aplicación lineal F.

PROBLEMAS RESUELTOS

REPRESENTACIONES MATRICIALES DE OPERADORES LINEALES

10.1. Supóngase que $F:\mathbf{R}^2 \to \mathbf{R}^2$ se define por $F(x, y) = (2y, 3x - y)$. Hallar la representación matricial de F relativa a la base usual $E = \{e_1 = (1, 0), e_2 = (0, 1)\}$.

Nótese primero que si $(a, b) \in \mathbf{R}^2$, entonces $(a, b) = ae_1 + be_2$:

$$\begin{aligned} F(e_1) &= F(1, 0) = (0, 3) \quad\;\; = 0e_1 + 3e_2 \\ F(e_2) &= F(0, 1) = (2, -1) = 2e_1 - \;\, e_2 \end{aligned} \qquad \text{y} \qquad [F]_E = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$$

Se ve que las filas de $[F]_E$ vienen dadas directamente por los coeficientes en las componentes de $F(x, y)$. Esto se generaliza a cualquier espacio K^n.

10.2. Encontrar la representación matricial del operador lineal F del Problema 10.1 relativa a la base $S = \{u_1 = (1, 3), u_2 = (2, 5)\}$.

Empezamos hallando las coordenadas de un vector arbitrario $(a, b) \in \mathbf{R}^2$ respecto a la base S. Tenemos

$$\begin{pmatrix} a \\ b \end{pmatrix} = x\begin{pmatrix} 1 \\ 3 \end{pmatrix} + y\begin{pmatrix} 2 \\ 5 \end{pmatrix} \qquad \text{o} \qquad \begin{aligned} x + 2y &= a \\ 3x + 5y &= b \end{aligned}$$

Despejamos x e y en términos de a y b consiguiendo $x = 2b - 5a$ e $y = 3a - b$. Así

$$(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$$

Tenemos $F(x, y) = (2y, 3x - y)$. De aquí

$$\begin{aligned} F(u_1) &= F(1, 3) = (6, 0) \quad\; = -30u_1 + 18u_2 \\ F(u_2) &= F(2, 5) = (10, 1) = -48u_1 + 29u_2 \end{aligned} \qquad \text{y} \qquad [F]_S = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$$

(**Nota:** Subrayamos que los coeficientes de u_1 y u_2 se escriben como columnas, no como filas, en cada representación matricial.)

10.3. Sea G el operador lineal en $\mathbf{R}^3$ definido por $G(x, y, z) = (2y + z, x - 4y, 3x)$.

a) Hallar la representación matricial de G relativa a la base

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\}$$

b) Comprobar que $[G][v] = [G(v)]$ para cualquier vector $v \in \mathbf{R}^3$.

Primero hallamos las coordenadas de un vector arbitrario $(a, b, c) \in \mathbf{R}^3$ respecto a la base S. Expresamos (a, b, c) como combinación lineal de w_1, w_2, w_3 empleando escalares desconocidos x, y y z:

$$(a, b, c) = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) = (x + y + z, x + y, x)$$

Igualamos entre sí las componentes correspondientes para obtener el sistema de ecuaciones

$$x + y + z = a \qquad x + y = b \qquad x = c$$

Resolvemos el sistema para x, y y z en términos de a, b y c encontrando $x = c$, $y = b - c$, $z = a - b$. Siendo así,

$$(a, b, c) = cw_1 + (b - c)w_2 + (a - b)w_3 \quad \text{o, equivalentemente,} \quad [(a, b, c)] = [c, b - c, a - b]^T$$

a) Puesto que $G(x, y, z) = (2y + z, x - 4y, 3x)$,

$$G(w_1) = G(1, 1, 1) = (3, -3, 3) = 3w_1 - 6w_2 + 6w_3$$
$$G(w_2) = G(1, 1, 0) = (2, -3, 3) = 3w_1 - 6w_2 + 5w_3$$
$$G(w_3) = G(1, 0, 0) = (0, 1, 3) = 3w_1 - 2w_2 - w_3$$

Escribimos las coordenadas de $G(w_1)$, $G(w_2)$, $G(w_3)$ como columnas llegando a

$$[G] = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$$

b) Expresamos $G(v)$ como combinación lineal de w_1, w_2, w_3, donde $v = (a, b, c)$ es un vector arbitrario en $\mathbf{R}^3$:

$$G(v) = G(a, b, c) = (2b + c, a - 4b, 3a) = 3aw_1 + (-2a - 4b)w_2 + (-a + 6b + c)w_3$$

o, equivalentemente,

$$[G(v)] = [3a, -2a - 4b, -a + 6b + c]^T$$

En consecuencia,

$$[G][v] = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix} \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix} = \begin{pmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{pmatrix} = [G(v)]$$

10.4. Sean $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y T el operador lineal en $\mathbf{R}^2$ definido por $T(v) = Av$ (donde v se escribe como vector columna). Encontrar la matriz de T en cada una de las siguientes bases:

a) $E = \{e_1 = (1, 0), e_2 = (0, 1)\}$, es decir, la base usual; *b*) $S = \{u_1 = (1, 3), u_2 = (2, 5)\}$.

a)

$$\left.\begin{aligned} T(e_1) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 1e_1 + 3e_2 \\ T(e_2) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2e_1 + 4e_2 \end{aligned}\right\} \quad \text{y así} \quad [T]_E = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Obsérvese que la matriz T en la base usual es precisamente la matriz original A que definía T. Véase Ejemplo 10.1 *c*).

b) Por el Problema 10.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Por tanto,

$$\left.\begin{array}{l} T(u_1) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 15 \end{pmatrix} = -5u_1 + 6u_2 \\ \\ T(u_2) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 12 \\ 26 \end{pmatrix} = -8u_1 + 10u_2 \end{array}\right\} \quad \text{y así} \quad [T]_S = \begin{pmatrix} -5 & -8 \\ 6 & 10 \end{pmatrix}$$

10.5. Cada uno de los conjuntos *a*) $\{1, t, e^t, te^t\}$ y *b*) $\{e^{3t}, te^{3t}, t^2e^{3t}\}$ es una base de un espacio vectorial V de funciones $f: \mathbf{R} \to \mathbf{R}$. Sea $\mathbf{D}$ el operador de derivación en V, esto es, $\mathbf{D}(f) = df/dt$. Hallar la matriz de $\mathbf{D}$ en las bases dadas.

a)

$$\begin{array}{ll} \mathbf{D}(1) = 0 & = 0(1) + 0(t) + 0(e^t) + 0(te^t) \\ \mathbf{D}(t) = 1 & = 1(1) + 0(t) + 0(e^t) + 0(te^t) \\ \mathbf{D}(e^t) = e^t & = 0(1) + 0(t) + 1(e^t) + 0(te^t) \\ \mathbf{D}(te^t) = e^t + te^t & = 0(1) + 0(t) + 1(e^t) + 1(te^t) \end{array} \quad \text{y así} \quad [\mathbf{D}] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

b)

$$\begin{array}{ll} \mathbf{D}(e^{3t}) = 3e^{3t} & = 3(e^{3t}) + 0(te^{3t}) + 0(t^2e^{3t}) \\ \mathbf{D}(te^{3t}) = e^{3t} + 3te^{3t} & = 1(e^{3t}) + 3(te^{3t}) + 0(t^2e^{3t}) \\ \mathbf{D}(t^2e^{3t}) = 2te^{3t} + 3t^2e^{3t} & = 0(e^{3t}) + 2(te^{3t}) + 3(t^2e^{3t}) \end{array} \quad \text{y así} \quad [\mathbf{D}] = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

10.6. Sea V el espacio vectorial de las matrices 2×2 con la base usual

$$\left\{E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

Sean $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y T el operador lineal en V definido por $T(A) = MA$. Hallar la representación matricial de T relativa a la base usual de V precedente.

Tenemos

$$T(E_1) = ME_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} = 1E_1 + 0E_2 + 3E_3 + 0E_4$$

$$T(E_2) = ME_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} = 0E_1 + 1E_2 + 0E_3 + 3E_4$$

$$T(E_3) = ME_3 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} = 2E_1 + 0E_2 + 4E_3 + 0E_4$$

$$T(E_4) = ME_4 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix} = 0E_1 + 2E_2 + 0E_3 + 4E_4$$

Por consiguiente,

$$[T] = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$$

(Como dim $V = 4$, toda representación matricial de un operador lineal en V debe ser una matriz 4-cuadrada.)

10.7. Considérese la base $S = \{(1, 0), (1, 1)\}$ de $\mathbf{R}^2$. Defínase $L: \mathbf{R}^2 \to \mathbf{R}^2$ por $L(1, 0) = (6, 4)$ y $L(1, 1) = (1, 5)$. (Recuérdese que una aplicación lineal está completamente definida por su acción sobre una base.) Hallar la representación matricial de L respecto a la base S.

Expresamos (6, 4) y después (1, 5) como combinaciones lineales de los vectores de la base consiguiendo

$$\begin{aligned} L(1, 0) &= (6, 4) = 2(1, 0) + 4(1, 1) \\ L(1, 1) &= (1, 5) = -4(1, 0) + 5(1, 1) \end{aligned} \qquad \text{y así} \qquad [L] = \begin{pmatrix} 2 & -4 \\ 4 & 5 \end{pmatrix}$$

10.8. Considérese la base usual $E = \{e_1, e_2, \ldots, e_n\}$ de K^n. Defínase $L: K^n \to K^n$ por $L(e_i) = v_i$. Mostrar que la matriz A que representa L relativa a la base usual E se obtiene escribiendo los vectores imagen $v_1, v_2, \ldots, v_n$ como columnas.

Supongamos $v_i = (a_{i1}, a_{i2}, \ldots, a_{in})$. Entonces $L(e_i) = v_i = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n$. De este modo,

$$[L] = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

como se pretendía.

10.9. Para cada uno de los siguientes operadores lineales L en $\mathbf{R}^2$, encontrar la matriz A que representa a L (respecto a la base usual de $\mathbf{R}^2$):

a) L se define por $L(1, 0) = (2, 4)$ y $L(0, 1) = (5, 8)$.

b) L es la rotación de 90° en el sentido contrario al de las agujas de un reloj, en $\mathbf{R}^2$.

c) L es la reflexión en la recta $y = -x$, en $\mathbf{R}^2$.

a) Dado que (1, 0) y (0, 1) forman la base usual en $\mathbf{R}^2$, escribimos sus imágenes como columnas (Problema 10.8) para conseguir

$$A = \begin{pmatrix} 2 & 5 \\ 4 & 8 \end{pmatrix}$$

b) Bajo la rotación L tenemos $L(1, 0) = (0, 1)$ y $L(0, 1) = (-1, 0)$. Así pues, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

c) Bajo la reflexión L tenemos $L(1, 0) = (0, -1)$ y $L(0, 1) = (-1, 0)$. Así pues, $A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$.

10.10. Demostrar el Teorema 10.1.

Supongamos, para $i = 1, \ldots, n$,

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = \sum_{j=1}^{n} a_{ij}u_j$$

Entonces $[T]_S$ es la matriz n-cuadrada cuya fila j-ésima es

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \qquad [1]$$

Supongamos ahora

$$v = k_1u_1 + k_2u_2 + \cdots + k_nu_n = \sum_{i=1}^{n} k_iu_i$$

Expresando un vector columna como la traspuesta de un vector fila,

$$[v]_S = (k_1, k_2, \ldots, k_n)^T \qquad [2]$$

Además, haciendo uso de la linealidad de T,

$$T(v) = T\left(\sum_{i=1}^{n} k_iu_i\right) = \sum_{i=1}^{n} k_i T(u_i) = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}u_j\right) =$$

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)u_j = \sum_{j=1}^{n}(a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)u_j$$

De este modo, $[T(v)]_S$ es el vector columna cuya entrada j-ésima es

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \qquad [3]$$

Por otra parte, la entrada j-ésima de $[T]_S[v]_S$ se obtiene multiplicando la fila j-ésima de $[T]_S$ por $[v]_S$, es decir, [1] por [2]. Pero el producto de [1] y [2] es [3], luego $[T]_S[v]_S$ y $[T(v)]_S$ tienen las mismas entradas. Siendo así, $[T]_S[v]_S = [T(v)]_S$.

10.11. Demostrar el Teorema 10.2.

Supongamos, para $i = 1, \ldots, n$,

$$F(u_i) = \sum_{j=1}^{n} a_{ij}u_j \quad \text{y} \quad G(u_i) = \sum_{j=1}^{n} b_{ij}u_j$$

Consideremos las matrices $A = (a_{ij})$ y $B = (b_{ij})$. Entonces $[F] = A^T$ y $[G] = B^T$. Tenemos, para $i = 1, \ldots, n$,

$$(F + G)(u_i) = F(u_i) + G(u_i) = \sum_{j=1}^{n}(a_{ij} + b_{ij})u_j$$

Siendo $A + B$ la matriz $(a_{ij} + b_{ij})$,

$$[A + B] = (A + B)^T = A^T + B^T = [A] + [B]$$

Asimismo, para $i = 1, \ldots, n$,

$$(kF)(u_i) = kF(u_i) = k\sum_{j=1}^{n} a_{ij}u_j = \sum_{j=1}^{n}(ka_{ij})u_j$$

Como kA es la matriz (ka_{ij}),

$$[kF] = (kA)^T = kA^T = k[F]$$

Finalmente, m es inyectiva puesto que una aplicación lineal está completamente determinada por su valores en una base, y es suprayectiva porque cada matriz $A = (a_{ij})$ en M es la imagen del operador lineal

$$F(u_i) = \sum_{j=1}^{n} a_{ij}u_j \qquad i = 1, \ldots, n$$

Queda, pues, demostrado el teorema.

10.12. Demostrar el Teorema 10.3.

Empleando la notación del Problema 10.11,

$$(G \circ F)(u_i) = G(F(u_i)) = G\left(\sum_{j=1}^{n} a_{ij} u_j\right) = \sum_{j=1}^{n} a_{ij} G(u_j) =$$
$$= \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk} u_k\right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) u_k$$

Recuérdese que AB es la matriz $AB = (c_{ik})$, donde $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$. De acuerdo con ello,

$$[G \circ F] = (AB)^T = B^T A^T = [G][F]$$

Queda, pues, demostrado el teorema.

10.13. Sea A una representación matricial de un operador T. Probar que $f(A)$ es la representación matricial de $f(T)$ para todo polinomio $f(t)$. [De esta manera, $f(T) = 0$ si y sólo si $f(A) = 0$.]

Sea ϕ la aplicación $T \mapsto A$, o sea, la que envía el operador T a su representación matricial A. Debemos probar que $\phi(f(T)) = f(A)$. Supongamos $f(t) = a_n t^n + \cdots + a_1 t + a_0$. Se demuestra por inducción en n el grado de $f(t)$.

Supongamos $n = 0$. Recuérdese que $\phi(I') = I$, siendo I' la aplicación identidad e I la matriz identidad. Así pues,

$$\phi(f(T)) = \phi(a_0 I') = a_0 \phi(I') = a_0 I = f(A)$$

de modo que el teorema es válido para $n = 0$.

Aceptemos ahora que el teorema es válido para polinomios de grado menor que n. En tal caso, por ser ϕ un isomorfismo entre álgebras,

$$\phi(f(T)) = \phi(a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') =$$
$$= a_n \phi(T)\phi(T^{n-1}) + \phi(a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') =$$
$$= a_n A A^{n-1} + (a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I) = f(A)$$

y queda demostrado el teorema.

10.14. Sea V el espacio vectorial de las funciones que tiene (sen θ, cos θ) como base y sea $\mathbf{D}$ el operador de derivación en V. Mostrar que $\mathbf{D}$ es un cero de $f(t) = t^2 + 1$.

Aplicamos $f(\mathbf{D})$ a cada vector de la base:

$$f(\mathbf{D})(\text{sen } \theta) = (\mathbf{D}^2 + I)(\text{sen } \theta) = \mathbf{D}^2(\text{sen } \theta) + I(\text{sen } \theta) = -\text{sen } \theta + \text{sen } \theta = 0$$
$$f(\mathbf{D})(\cos \theta) = (\mathbf{D}^2 + I)(\cos \theta) = \mathbf{D}^2(\cos \theta) + I(\cos \theta) = -\cos \theta + \cos \theta = 0$$

Como cada vector de la base se aplica en 0, todo vector $v \in V$ será aplicado también en 0 por $f(\mathbf{D})$. Así $f(\mathbf{D}) = \mathbf{0}$.

[Este resultado era de esperar ya que, según el Ejemplo 10.5 *a*), $f(t)$ es el polinomio característico de $\mathbf{D}$.]

CAMBIO DE BASE. MATRICES SIMILARES

10.15. Defínase $F: \mathbf{R}^2 \to \mathbf{R}^2$ como $F(x, y) = (4x - y, 2x + y)$ y considérense las bases de $\mathbf{R}^2$:

$$E = \{e_1 = (1, 0), e_2 = (0, 1)\} \qquad \text{y} \qquad S = \{u_1 = (1, 3), u_2 = (2, 5)\}$$

a) Hallar la matriz de cambio de base P de E a S, la matriz de cambio de base Q de S a E y comprobar que $Q = P^{-1}$.

b) Hallar la matriz A que representa F en la base E, la matriz B que representa F en la base S y comprobar que $B = P^{-1}AP$.

c) Hallar la traza tr F, el determinante det (F) y el polinomio característico $\Delta(t)$ de F.

a) Dado que E es la base usual, escribimos los elementos de S como columnas y obtenemos la matriz de cambio de base

$$P = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

Resolviendo el sistema $u_1 = e_1 + 3e_2$, $u_2 = 2e_1 + 5e_2$, para e_1 y e_2,

$$\begin{aligned} e_1 &= -5u_1 + 3u_2 \\ e_2 &= 2u_1 - u_2 \end{aligned} \qquad \text{y así} \qquad Q = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$

Tenemos

$$PQ = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

b) Escribimos los coeficientes de x e y como filas (Problema 10.3) para llegar a

$$A = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$$

Como $F(x, y) = (4x - y, 2x + y)$ y $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$,

$$\begin{aligned} F(u_1) &= F(1, 3) = (1, 5) = 5u_1 - 2u_2 \\ F(u_2) &= F(2, 5) = (3, 9) = 3u_1 \end{aligned} \qquad \text{y así} \qquad B = \begin{pmatrix} 5 & 3 \\ -2 & 0 \end{pmatrix}$$

Tenemos

$$P^{-1}AP = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ -2 & 0 \end{pmatrix} = B$$

c) Usamos A (o B) obteniendo

$$\text{tr } F = \text{tr } A = 4 + 1 = 5 \qquad \det(F) = \det(A) = 4 + 2 = 6$$
$$\Delta(t) = t^2 - (\text{tr } A)t + \det(A) = t^2 - 5t + 6$$

10.16. Sea G el operador lineal en $\mathbf{R}^3$ definido por $G(x, y, z) = (2y + z, x - 4y, 3x)$ y considérense la base usual E de $\mathbf{R}^3$ y las bases S de $\mathbf{R}^3$:

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\}$$

a) Encontrar la matriz de cambio de base P de E a S, la matriz de cambio de base Q de S a E y comprobar que $Q = P^{-1}$.

b) Comprobar que $[G]_S = P^{-1}[G]_E P$.

c) Hallar la traza, determinante y polinomio característico de G.

a) Siendo E la base usual, escribimos los elementos de S como columnas y conseguimos la matriz de cambio de base

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Mediante el familiar proceso de inversión (véase el Problema 10.3) obtenemos

$$\begin{aligned} e_1 &= 0w_1 + 0w_2 + 1w_3 \\ e_2 &= 0w_1 + 1w_2 - 1w_3 \\ e_3 &= 1w_1 - 1w_2 + 0w_3 \end{aligned} \qquad \text{y así} \qquad Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

Tenemos

$$PQ = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

b) De los Problemas 10.1 y 10.3, $[G]_E = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$ y $[G]_S = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$. De este modo,

$$P^{-1}[G]_E P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix} = [G]_S$$

c) Utilizamos $[G]_E$ (la matriz más simple) para llegar a

$$\operatorname{tr} G = 0 - 4 + 0 = -4 \qquad \det(G) = 12 \qquad \text{e} \qquad \Delta(t) = t^3 + 4t^2 - 5t - 12$$

10.17. Hallar la traza y el determinante del siguiente operador en $\mathbf{R}^3$:

$$T(x, y, z) = (a_1x + a_2y + a_3z,\ b_1x + b_2y + b_3z,\ c_1x + c_2y + c_3z)$$

Empezamos por encontrar una representación matricial A de T. Eligiendo la base usual E,

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

En tal caso,

$$\operatorname{tr} T = \operatorname{tr} A = a_1 + b_2 + c_3$$

y

$$\det(T) = \det(A) = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_bc_3 - a_2b_1c_3 - a_1b_3c_2$$

10.18. Sean V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$ y $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Sea T el operador lineal en V definido mediante $T(A) = MA$. Hallar la traza y el determinante de T.

Debemos encontrar primero una representación matricial de T. Escogemos la base usual de V para llegar (Problema 10.6) a la representación matricial:

$$[T] = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$$

Entonces $\operatorname{tr} T = 1 + 1 + 4 + 4 = 10$ y $\det(T) = 4$.

10.19. Demostrar el Teorema 10.4.

Sea v cualquier vector en V. Entonces, en virtud del Teorema 5.27, $P[v]_{S'} = [v]_S$. Por consiguiente,

$$P^{-1}[T]_S P[v]_{S'} = P^{-1}[T]_S[v]_S = P^{-1}[T(v)]_S = [T(v)]_{S'}$$

Pero $[T]_{S'}[v]_{S'} = [T(v)]_{S'}$, luego

$$P^{-1}[T]_S P[v]_{S'} = [T]_{S'}[v]_{S'}$$

Como la aplicación $v \mapsto [v]_{S'}$ es sobre K^n, tenemos $P^{-1}[T]_S PX = [T]_{S'}X$ para todo $X \in K^n$. De este modo, $P^{-1}[T]_S P = [T]_{S'}$, como se pretendía.

DIAGONALIZACION DE OPERADORES LINEALES. VALORES PROPIOS Y VECTORES PROPIOS

10.20. Hallar los valores propios y vectores propios linealmente independientes del siguiente operador lineal en $\mathbf{R}^2$ y, si es diagonalizable, encontrar una representación diagonal D: $F(x, y) = (6x - y, 3x + 2y)$.

Comenzamos hallando la matriz A que representa F en la base usual de $\mathbf{R}^2$ escribiendo los coeficientes de x e y como filas:

$$A = \begin{pmatrix} 6 & -1 \\ 3 & 2 \end{pmatrix}$$

El polinomio característico $\Delta(t)$ de F es, pues,

$$\Delta(t) = t^2 - (\operatorname{tr} A)t + |A| = t^2 - 8t + 15 = (t-3)(t-5)$$

Siendo así, $\lambda_1 = 3$ y $\lambda_2 = 5$ son valores propios de F. Encontramos los correspondientes vectores propios como sigue:

i) Restamos $\lambda_1 = 3$ a los elementos diagonales de A para obtener la matriz $M = \begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix}$ que corresponde al sistema homogéneo $3x - y = 0$. Aquí $v_1 = (1, 3)$ es una solución no nula y por tanto un vector propio de F perteneciente a $\lambda_1 = 3$.

ii) Restamos $\lambda_2 = 5$ a los elementos diagonales de A llegando a $M = \begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix}$ correspondiente al sistema homogéneo $x - y = 0$. Aquí $v_2 = (1, 1)$ es una solución no nula y por ende un vector propio de F perteneciente a $\lambda_2 = 5$.

Entonces $S = \{v_1, v_2\}$ es una base de $\mathbf{R}^2$ que consiste en vectores propios de F. Así F es diagonalizable y tiene la representación matricial $D = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$.

10.21. Sea L el operador lineal en $\mathbf{R}^2$ que refleja los puntos en la recta $y = kx$ (con $k \neq 0$). Véase la Figura 10-2.

a) Probar que $v_1 = (k, 1)$ y $v_2 = (1, -k)$ son vectores propios de L.

b) Probar que L es diagonalizable y hallar una representación diagonal D.

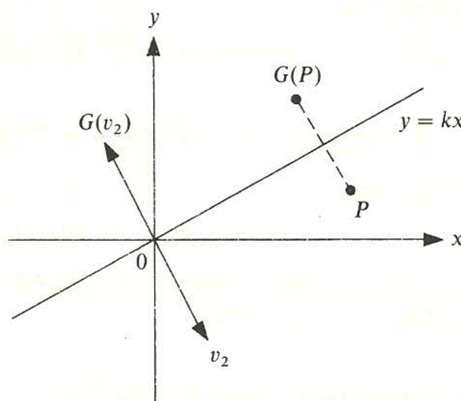


Figura 10-2.

a) El vector $v_1 = (k, 1)$ yace en la recta $y = kx$ y por consiguiente L lo deja fijo, esto es, $L(v_1) = v_1$. De este modo, v_1 es un vector propio de L perteneciente al valor propio $\lambda_1 = 1$. El vector $v_2 = (1, -k)$ es perpendicular a la recta $y = kx$, luego L refleja v_2 en su opuesto, es decir, $L(v_2) = -v_2$. Así v_2 es un vector propio de L perteneciente al valor propio $\lambda_2 = -1$.

b) Aquí $S = \{v_1, v_2\}$ es una base de $\mathbf{R}^2$ que consiste en vectores propios de L. Por tanto, L es diagonalizable y tiene la representación diagonal (relativa a S) $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

10.22. Hallar todos los valores propios y una base para cada espacio propio del operador $T: \mathbf{R}^3 \to \mathbf{R}^3$ definido por $T(x, y, z) = (2x + y, y - z, 2y + 4z)$. ¿Es diagonalizable T? En caso afirmativo, hallar una representación diagonal D.

Primero encontramos la matriz A que representa T en la base usual de $\mathbf{R}^3$ escribiendo los coeficientes de x, y, z como filas:

$$A = [T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

El polinomio característico $\Delta(t)$ de T es, pues,

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{vmatrix} = (t-2)^2(t-3)$$

Siendo así, 2 y 3 son los valores propios de T.

Hallemos una base del espacio propio E_2 del valor propio 2. Sustituimos $t = 2$ en $tI - A$ para llegar al sistema homogéneo.

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} -y = 0 \\ y + z = 0 \\ -2y - 2z = 0 \end{cases} \quad \text{o} \quad \begin{cases} y = 0 \\ y + z = 0 \end{cases}$$

El sistema sólo tiene una solución independiente, por ejemplo, $x = 1$, $y = 0$, $z = 0$. De esta manera, $u = (1, 0, 0)$ constituye una base del espacio propio E_2.

Hallemos una base del espacio propio E_3 del valor propio 3. Sustituimos $t = 3$ en $tI - A$ para llegar al sistema homogéneo

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{o} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \\ -2y - z = 0 \end{cases} \quad \text{o} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \end{cases}$$

El sistema sólo tiene una solución independiente, por ejemplo, $x = 1$, $y = 1$, $z = -2$. Así $v = (1, 1, -2)$ forma una base del espacio propio E_3.

Obsérvese que T no es diagonalizable, porque sólo tiene dos vectores propios linealmente independientes.

10.23. Mostrar que 0 es un valor propio de T si y sólo si T es singular.

Tenemos que 0 es un valor propio de T si y sólo si existe un vector no nulo v tal que $T(v) = 0v = 0$, o sea, si y sólo si T es singular.

10.24. Supóngase que λ es un valor propio de un operador invertible T. Probar que λ^{-1} es un valor propio de T^{-1}.

Como T es invertible, es también no singular, luego, de acuerdo con el Problema 10.23, $\lambda \neq 0$.

Por definición de valor propio, existe un vector no nulo v para el que $T(v) = \lambda v$. Aplicando T^{-1} a ambos miembros obtenemos $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$. De aquí $T^{-1}(v) = \lambda^{-1}v$, esto es λ^{-1} es un valor propio de T^{-1}.

10.25. Supóngase que $\dim V = n$. Sea $T: V \to V$ un operador invertible. Probar que T^{-1} es igual a un polinomio en T cuyo grado no excede n.

Sea $m(t)$ el polinomio mínimo de T. Entonces $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$, donde $r \leqslant n$. Dado que T es invertible, $a_0 \neq 0$. Tenemos

$$m(T) = T^r + a_{r-1}T^{r-1} + \cdots + a_1 T + a_0 I = 0$$

Por tanto,

$$-\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)T = I \qquad \text{y} \qquad T^{-1} = -\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)$$

10.26. Demostrar el Teorema 10.7.

Se demuestra por inducción en n. Si $n = 1$, u_1 es linealmente independiente puesto que $u_1 \neq 0$. Supongamos que $n > 1$ y que

$$a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = 0 \qquad [1]$$

donde los a_i son escalares. Aplicando T a la relación anterior llegamos, por linealidad, a

$$a_1 T(u_1) + a_2 T(u_2) + \cdots + a_n T(u_n) = T(0) = 0$$

Pero, por hipótesis, $T(u_i) = \lambda_i u_i$, luego

$$a_1 \lambda_1 u_1 + a_2 \lambda_2 u_2 + \cdots + a_n \lambda_n u_n = 0 \qquad [2]$$

Por otra parte, multiplicando [1] por λ_n,

$$a_1 \lambda_n u_1 + a_2 \lambda_n u_2 + \cdots + a_n \lambda_n u_n = 0 \qquad [3]$$

Restando ahora [3] a [2],

$$a_1(\lambda_1 - \lambda_n)u_1 + a_2(\lambda_2 - \lambda_n)u_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)u_{n-1} = 0$$

Por inducción, $u_1, u_2, \ldots, u_{n-1}$ son linealmente independientes, luego cada uno de los coeficientes precedentes es 0. Como los λ_i son distintos, $\lambda_i - \lambda_n \neq 0$ para $i \neq n$. De aquí $a_1 = \cdots = a_{n-1} = 0$. Sustituyendo esto en [1] conseguimos $a_n u_n = 0$ y por ende $a_n = 0$. Los u_i son, pues, linealmente independientes.

10.27. Demostrar el Teorema 10.10.

Supongamos que la multiplicidad geométrica de λ es r. En ese caso hay r vectores propios linealmente independientes $v_1, \ldots, v_r$ pertenecientes a λ. Extendemos el conjunto $\{v_i\}$ a una base de $V: \{v_1, \ldots, v_r, w_1, \ldots, w_s\}$. Tenemos

$$\begin{aligned}
T(v_1) &= \lambda v_1 \\
T(v_2) &= \quad \lambda v_2 \\
&\cdots\cdots\cdots\cdots\cdots \\
T(v_r) &= \qquad \lambda v_r \\
T(w_1) &= a_{11}v_1 + \cdots + a_{1r}v_r + b_{11}w_1 + \cdots + b_{1s}w_s \\
T(w_2) &= a_{21}v_1 + \cdots + a_{2r}v_r + b_{21}w_1 + \cdots + b_{2s}w_s \\
&\cdots\cdots\cdots\cdots\cdots \\
T(w_s) &= a_{s1}v_1 + \cdots + a_{sr}v_r + b_{s1}w_1 + \cdots + b_{ss}w_s
\end{aligned}$$

La matriz de T en la base anterior es

$$M = \left(\begin{array}{cccc|cccc}
\lambda & 0 & \ldots & 0 & a_{11} & a_{21} & \cdots & a_{s1} \\
0 & \lambda & \ldots & 0 & a_{12} & a_{22} & \cdots & a_{s2} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & \ldots & \lambda & a_{1r} & a_{2r} & \cdots & a_{sr} \\
\hline
0 & 0 & \ldots & 0 & b_{11} & b_{21} & \cdots & b_{r1} \\
0 & 0 & \ldots & 0 & b_{12} & b_{22} & \cdots & b_{r2} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & \ldots & 0 & b_{1s} & b_{2s} & \cdots & b_{ss}
\end{array}\right) = \left(\begin{array}{c|c} \lambda I_r & A \\ \hline 0 & B \end{array}\right)$$

donde $A = (a_{ij})^T$ y $B = (b_{ij})^T$.

Siendo M una matriz triangular por bloques, el polinomio característico de λI_r, que es $(t-\lambda)^r$, debe ser divisor del de M y por consiguiente del de T. De este modo, la multiplicidad algebraica de λ para el operador T es al menos r, como se pedía.

10.28. Sea $\{v_1, \ldots, v_n\}$ una base de V. Sea $T: V \to V$ un operador para el cual $T(v_1) = 0$, $T(v_2) = a_{21}v_1$, $T(v_3) = a_{31}v_1 + a_{32}v_2$, ..., $T(v_n) = a_{n1}v_1 + \cdots + a_{n,n-1}v_{n-1}$. Probar que $T^n = \mathbf{0}$.

Basta probar que

$$T^j(v_j) = 0 \qquad (*)$$

para $j = 1, \ldots, n$. La razón es que entonces se deriva que

$$T^n(v_j) = T^{n-j}(T^j(v_j)) = T^{n-j}(0) = 0 \qquad \text{para } j = 1, \ldots, n$$

y al ser $\{v_1, \ldots, v_n\}$ una base, $T^n = \mathbf{0}$.

Demostramos (*) por inducción en j. El caso $j = 1$ es cierto por hipótesis. El paso inductivo se sigue (para $j = 2, \ldots, n$) de

$$\begin{aligned} T^j(v_j) &= T^{j-1}(T(v_j)) = T^{j-1}(a_{j1}v_1 + \cdots + a_{j,j-1}v_{j-1}) = \\ &= a_{j1}T^{j-1}(v_1) + \cdots + a_{j,j-1}T^{j-1}(v_{j-1}) = \\ &= a_{j1}0 + \cdots + a_{j,j-1}0 = 0 \end{aligned}$$

Nota: Obsérvese que la representación matricial de T en la base precedente es triangular con elementos diagonales 0:

$$\begin{pmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

REPRESENTACIONES MATRICIALES DE APLICACIONES LINEALES

10.29. Sea $F: \mathbf{R}^3 \to \mathbf{R}^2$ la aplicación lineal definida por $F(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$.

a) Hallar la matriz de F en las siguientes bases de $\mathbf{R}^3$ y $\mathbf{R}^2$:

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\} \qquad S' = \{u_1 = (1, 3), u_2 = (2, 5)\}$$

b) Comprobar que la acción de F es preservada por su representación matricial; esto es, para todo $v \in \mathbf{R}^3$, $[F]_S^{S'}[v]_S = [F(v)]_{S'}$

a) Del Problema 10.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Siendo así,

$$\begin{aligned} F(w_1) &= F(1, 1, 1) = (1, -1) = -7u_1 + 4u_2 \\ F(w_2) &= F(1, 1, 0) = (5, -4) = -33u_1 + 19u_2 \\ F(w_3) &= F(1, 0, 0) = (3, 1) = -13u_1 + 8u_2 \end{aligned}$$

Escribimos las coordenadas de $F(w_1)$, $F(w_2)$, $F(w_3)$ como columnas para conseguir

$$[F]_S^{S'} = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

b) Si $v = (x, y, z)$, según el Problema 10.3, $v = zw_1 + (y - z)w_2 + (x - y)w_3$. Asimismo,

$$F(v) = (3x + 2y - 4z, x - 5y + 3z) = (-13x - 20y + 26z)u_1 + (8x + 11y - 15z)u_2$$

Por tanto,

$$[v]_S = (z, y - z, x - y)^T \qquad \text{y} \qquad [F(v)]_{S'} = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix}$$

De este modo,

$$[F]_S^{S'}[v]_S = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}\begin{pmatrix} z \\ y - z \\ x - y \end{pmatrix} = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix} = [F(v)]_{S'}$$

10.30. Sea $F: K^n \to K^m$ la aplicación lineal definida por

$$F(x_1, x_2, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, a_{21}x_1 + \cdots + a_{2n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n)$$

Mostrar que la representación matricial de F relativa a las bases usuales de K^n y K^m viene dada por

$$[F] = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

Es decir, las filas de $[F]$ se obtienen de los coeficientes de las x_i en las componentes de $F(x_1, \ldots, x_n)$.

Tenemos

$$\begin{aligned} F(1, 0, \ldots, 0) &= (a_{11}, a_{21}, \ldots, a_{m1}) \\ F(0, 1, \ldots, 0) &= (a_{12}, a_{22}, \ldots, a_{m2}) \\ &\ldots\ldots\ldots\ldots \\ F(0, 0, \ldots, 1) &= (a_{1n}, a_{2n}, \ldots, a_{mn}) \end{aligned} \qquad \text{y así} \qquad [F] = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

10.31. Encontrar la representación matricial de cada una de las aplicaciones lineales escritas a continuación respecto a las bases usuales de los $\mathbf{R}^n$:

$F: \mathbf{R}^2 \to \mathbf{R}^3$ definida como $F(x, y) = (3x - y, 2x + 4y, 5x - 6y)$

$F: \mathbf{R}^4 \to \mathbf{R}^2$ definida como $F(x, y, s, t) = (3x - 4y + 2s - 5t, 5x + 7y - s - 2t)$

$F: \mathbf{R}^3 \to \mathbf{R}^4$ definida como $F(x, y, z) = (2x + 3y - 8z, x + y + z, 4x - 5z, 6y)$

De acuerdo con el Problema 10.30, solamente es necesario mirar los coeficientes de las variables en $F(x, y, \ldots)$. Así pues,

$$[F] = \begin{pmatrix} 3 & -1 \\ 2 & 4 \\ 5 & -6 \end{pmatrix} \qquad [F] = \begin{pmatrix} 3 & -4 & 2 & -5 \\ 5 & 7 & -1 & -2 \end{pmatrix} \qquad [F] = \begin{pmatrix} 2 & 3 & -8 \\ 1 & 1 & 1 \\ 4 & 0 & -5 \\ 0 & 6 & 0 \end{pmatrix}$$

10.32. Sea la aplicación $T: \mathbf{R}^2 \to \mathbf{R}^2$ definida como $T(x, y) = (2x - 3y, x + 4y)$. Hallar la matriz de T relativa, respectivamente, a las siguientes bases de $\mathbf{R}^2$:

$$E = \{e_1 = (1, 0), e_2 = (0, 1)\} \qquad \text{y} \qquad S = \{u_1 = (1, 3), u_2 = (2, 5)\}$$

(Podemos ver T como una aplicación lineal de un espacio en otro, teniendo cada uno su propia base.)

Del Problema 10.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. Por tanto,

$$\begin{aligned} T(e_1) &= T(1, 0) = (2, 1) \quad = -8u_1 + 5u_2 \\ T(e_2) &= T(0, 1) = (-3, 4) = 23u_1 - 13u_2 \end{aligned} \qquad \text{y así} \qquad [T]_E^S = \begin{pmatrix} -8 & 23 \\ 5 & -13 \end{pmatrix}$$

10.33. Sea $A = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}$. Recuérdese que A determina una aplicacion lineal $F: \mathbf{R}^3 \to \mathbf{R}^2$ definida mediante $F(v) = Av$, donde v se escribe como vector columna. Encontrar la representación matricial de F relativa a las siguientes bases de $\mathbf{R}^3$ y $\mathbf{R}^2$.

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\} \qquad S' = \{u_1 = (1, 3), u_2 = (2, 5)\}$$

Del Problema 10.2, $(a, b) = (-5a + 2b)u_1 + (3a - b)u_2$. De esta manera,

$$F(w_1) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = -12u_1 + 8u_2$$

$$F(w_2) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \end{pmatrix} = -41u_1 + 24u_2$$

$$F(w_3) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -8u_1 + 5u_2$$

Escribir como columnas los coeficientes de $F(w_1)$, $F(w_2)$, $F(w_3)$ proporciona

$$[F]_S^{S'} = \begin{pmatrix} -12 & -41 & -8 \\ 8 & 24 & 5 \end{pmatrix}$$

10.34. Demostrar el Teorema 10.16.

Supongamos que dim $V = m$ y dim $U = n$. Sea W el núcleo de F y U', su imagen. Se nos dice que rango $F = r$, luego la dimensión del núcleo de F es $m - r$. Sea $\{w_1, \ldots, w_{m-r}\}$ una base del núcleo de F. Extendámosla a una base de V:

$$\{v_1, \ldots, v_r, w_1, \ldots, w_{m-r}\}$$

Tomemos
$$u_1 = F(v_1), u_2 = F(v_2), \ldots, u_r = F(v_r)$$

Hacemos notar que $\{u_1, \ldots, u_r\}$ es una base de U', la imagen de F. Extendemos ésta a una base

$$\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$$

de U. Obsérvese que

$$\begin{array}{lllll}
F(v_1) & = u_1 & = 1u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
F(v_2) & = u_2 & = 0u_1 + 1u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
\cdots & \cdots & \cdots \\
F(v_r) & = u_r & = 0u_1 + 0u_2 + \cdots + 1u_r + 0u_{r+1} + \cdots + 0u_n \\
F(w_1) & = 0 & = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
\cdots & \cdots & \cdots \\
F(w_{m-r}) & = 0 & = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n
\end{array}$$

Así pues, la matriz de F en las bases precedentes tiene la forma requerida.

PROBLEMAS SUPLEMENTARIOS

REPRESENTACIONES MATRICIALES DE OPERADORES LINEALES

10.35. Hallar la representación matricial de cada uno de los siguientes operadores lineales T en $\mathbf{R}^3$, relativa a la base usual:

a) $T(x, y, z) = (x, y, 0)$.

b) $T(x, y, z) = (2x - 7y - 4z, 3x + y + 4z, 6x - 8y + z)$.

c) $T(x, y, z) = (z, y + z, x + y + z)$.

10.36. Hallar la matriz de cada operador T del Problema 10.35 respecto a la base

$$S = \{u_1 = (1, 1, 0), u_2 = (1, 2, 3), u_3 = (1, 3, 5)\}$$

10.37. Sea $\mathbf{D}$ el operador de derivación, o sea, $\mathbf{D}(f) = df/dt$. Cada uno de los conjuntos que siguen es una base de un espacio vectorial V de funciones $f: \mathbf{R} \to \mathbf{R}$. Encontrar la matriz de $\mathbf{D}$ en cada base:

a) $\{e^t, e^{2t}, te^{2t}\}$, *b*) $\{\text{sen } t, \cos t\}$, *c*) $\{e^{5t}, te^{5t}, t^2e^{5t}\}$, *d*) $\{1, t, \text{sen } 3t, \cos 3t\}$.

10.38. Considérese el cuerpo complejo $\mathbf{C}$ como espacio vectorial sobre el cuerpo real $\mathbf{R}$. Sea T el operador de conjugación en $\mathbf{C}$, esto es, $T(z) = \bar{z}$. Hallar la matriz de T en cada base: *a*) $\{1, i\}$, *b*) $\{1 + i, 1 + 2i\}$.

10.39. Sean V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$ y $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Hallar la matriz de cada uno de los operadores lineales T en V escritos a continuación en la base usual (véase el Problema 10.18):

a) $T(A) = MA$, *b)* $T(A) = AM$, *c)* $T(A) = MA - AM$.

10.40. Denótense por $\mathbf{1}_V$ y $\mathbf{0}_V$ los operadores identidad y cero, respectivamente, en un espacio vectorial V. Demostrar que para cualquier base S de V: *a)* $[\mathbf{1}_V]_S = I$, la matriz identidad; *b)* $[\mathbf{0}_V]_S = 0$, la matriz cero.

CAMBIO DE BASE. MATRICES SIMILARES

10.41. Considérense las siguientes bases de $\mathbf{R}^2$: $E = \{e_1 = (1, 0), e_2 = (0, 1)\}$ y $S = \{u_1 = (1, 2), u_2 = (2, 3)\}$.

a) Encontrar las matrices de cambio de base P y Q desde E hasta S y desde S hasta E, respectivamente. Comprobar que $Q = P^{-1}$.

b) Mostrar que $[v]_E = P[v]_S$ para todo vector $v \in \mathbf{R}^2$.

c) Comprobar que $[T]_S = P^{-1}[T]_E P$ para el operador lineal $T(x, y) = (2x - 3y, x + y)$.

10.42. Hallar la traza y el determinante de cada una de las aplicaciones lineales en $\mathbf{R}^3$:

a) $F(x, y, z) = (x + 3y, 3x - 2z, x - 4y - 3z)$, *b)* $G(x, y, z) = (x + y - z, x + 3y, 4y + 3z)$.

10.43. Supóngase que $S = \{u_1, u_2\}$ es una base de V y $T: V \to V$ un operador lineal para el que $T(u_1) = 3u_1 - 2u_2$ y $T(u_2) = u_1 + 4u_2$. Supóngase, asimismo, que $S' = \{w_1, w_2\}$ es una base de V en la que $w_1 = u_1 + u_2$ y $w_2 = 2u_1 + 3u_2$. Hallar la matriz de T en la base S'.

10.44. Considérense las bases $\{1, i\}$ y $\{1 + i, 1 + 2i\}$ del espacio vectorial $\mathbf{C}$ sobre el cuerpo real $\mathbf{R}$. *a)* Encontrar las matrices de cambio de base P y Q desde S hasta S' y desde S' hasta S, respectivamente. Comprobar que $Q = P^{-1}$. *b)* Probar que $[T]_{S'} = P^{-1}[T]_S P$ para el operador de conjugación T del Problema 10.38.

DIAGONALIZACION DE OPERADORES LINEALES. VALORES PROPIOS Y VECTORES PROPIOS

10.45. Supóngase que v es un vector propio de dos operadores S y T. Mostrar que v es también un vector propio del operador $aS + bT$, donde a y b son escalares cualesquiera.

10.46. Supóngase que v es un vector propio de un operador T perteneciente al valor propio λ. Probar que, para $n > 0$, v es también un vector propio de T^n perteneciente a λ^n.

10.47. Supóngase que λ es un valor propio de un operador T y $f(t)$ un polinomio. Probar que $f(\lambda)$ es un valor propio de $f(T)$.

10.48. Mostrar que un operador lineal T es diagonalizable si y sólo si su polinomio mínimo es un producto de factores lineales distintos.

10.49. Sean S y T operadores lineales tales que $ST = TS$. Sean λ un valor propio de T y W su espacio propio. Probar que W es invariante bajo S, es decir, $S(W) \subseteq W$.

REPRESENTACIONES MATRICIALES DE APLICACIONES LINEALES

10.50. Hallar la representación matricial relativa a las bases usuales de los $\mathbf{R}^n$ de la aplicación lineal $F: \mathbf{R}^3 \to \mathbf{R}^2$ definida por $F(x, y, z) = (2x - 4y + 9z, 5x + 3y - 2z)$.

10.51. Sea $F: \mathbf{R}^3 \to \mathbf{R}^2$ la aplicación lineal definida por $F(x, y, z) = (2x + y - z, 3x - 2y + 4z)$. Encontrar la matriz de F en las siguientes bases de $\mathbf{R}^3$ y $\mathbf{R}^2$:

$$S = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\} \qquad S' = \{v_1 = (1, 3), v_2 = (1, 4)\}$$

Comprobar que, para cualquier vector $v \in \mathbf{R}^3$, $[F]_S^{S'}[v]_S = [F(v)]_{S'}$.

10.52. Sean S y S' bases de V y 1_V, la aplicación identidad en V. Probar que la matriz de 1_V relativa a las bases S y S' es la inversa de la matriz de cambio de base P desde S hasta S'; esto es, $[1_V]_S^{S'} = P^{-1}$.

10.53. Demostrar el Teorema 10.12.

10.54. Demostrar el Teorema 10.13.

10.55. Demostrar el Teorema 10.15.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

10.35. *a)* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, *b)* $\begin{pmatrix} 2 & -7 & -4 \\ 3 & 1 & 4 \\ 6 & -8 & 1 \end{pmatrix}$, *c)* $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

10.36. *a)* $\begin{pmatrix} 1 & 3 & 5 \\ 0 & -5 & -10 \\ 0 & 3 & 6 \end{pmatrix}$, *b)* $\begin{pmatrix} 15 & 51 & 104 \\ -49 & -191 & -351 \\ 29 & 116 & 208 \end{pmatrix}$, *c)* $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 2 & 3 \\ 1 & 0 & 0 \end{pmatrix}$

10.37. *a)* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, *b)* $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, *c)* $\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 2 \\ 0 & 0 & 5 \end{pmatrix}$, *d)* $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 3 & 0 \end{pmatrix}$

10.38. *a)* $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, *b)* $\begin{pmatrix} -3 & 4 \\ -2 & -3 \end{pmatrix}$

10.39. *a)* $\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix}$, *b)* $\begin{pmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & a & c \\ 0 & 0 & b & d \end{pmatrix}$, *c)* $\begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}$

10.41. $P = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $Q = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$

10.42. *a*) $-2, 13, t^3 + 2t^2 - 20t - 13$; *b*) $7, 2, t^3 - 7t^2 + 14t - 2$.

10.43. $\begin{pmatrix} 8 & 11 \\ -2 & -1 \end{pmatrix}$

10.44. $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $Q = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$

10.50. $\begin{pmatrix} 2 & -4 & 9 \\ 5 & 3 & -2 \end{pmatrix}$

10.51. $\begin{pmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{pmatrix}$

CAPITULO 11

Formas canónicas

11.1. INTRODUCCION

Sea T un operador lineal en un espacio vectorial de dimensión finita. Como se vio en el Capítulo 10, T puede no tener una representación matricial diagonal. No obstante, todavía es posible «simplificar» la representación matricial de T de numerosas formas. Este es el tema principal del capítulo. En particular, obtendremos el teorema de descomposición primaria y las formas canónicas triangular, de Jordan y racional.

Cabe comentar que las formas canónicas triangular y de Jordan para T existen si y sólo si el polinomio característico $\Delta(t)$ de T tiene todas sus raíces en el cuerpo base K. Esto siempre se verifica si K es el cuerpo complejo $\mathbf{C}$, pero puede no verificarse si K es el cuerpo real $\mathbf{R}$.

Introducimos también la idea de *espacio cociente*. Es una herramienta muy poderosa y se empleará en la demostración de la existencia de las formas canónicas triangular y racional.

11.2. FORMA TRIANGULAR

Sea T un operador lineal en un espacio vectorial n-dimensional V. Supongamos que T admite como representación la matriz triangular

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \cdots\cdots & \\ & & & a_{nn} \end{pmatrix}$$

Entonces el polinomio característico de T,

$$\Delta(t) = |tI - A| = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

es producto de factores lineales. El recíproco es cierto también y es un importante teorema; a saber (véase la demostración en el Problema 11.28),

Teorema 11.1: Sea $T: V \to V$ un operador lineal cuyo polinomio característico se factoriza en polinomio lineales. Existe una base de V en la que T se representa por una matriz triangular.

Teorema 11.1 (forma alternativa): Sea A una matriz cuadrada cuyo polinomio característico se factoriza en polinomios lineales. Entonces A es similar a una matriz triangular, es decir, existe una matriz invertible P tal que $P^{-1}AP$ es triangular.

Decimos que un operador T puede llevarse a forma triangular si se puede representar por una matriz triangular. Nótese que en este caso los valores propios de T son precisamente aquellas entradas situadas en la diagonal principal. Demos una aplicación de esta observación.

EJEMPLO 11.1. Sea A una matriz cuadrada sobre el cuerpo complejo $\mathbf{C}$. Supongamos que λ es un valor propio de A^2. Probemos que $\sqrt{\lambda}$ o $-\sqrt{\lambda}$ es un valor propio de A. Sabemos por el Teorema 11.1 que A es similar a una matriz triangular

$$B = \begin{pmatrix} \mu_1 & * & \ldots & * \\ & \mu_2 & \ldots & * \\ & & \ldots\ldots & \\ & & & \mu_n \end{pmatrix}$$

Por tanto, A^2 es similar a la matriz

$$B^2 = \begin{pmatrix} \mu_1^2 & * & \ldots & * \\ & \mu_2^2 & \ldots & * \\ & & \ldots\ldots & \\ & & & \mu_n^2 \end{pmatrix}$$

Dado que las matrices similares tienen los mismos valores propios, $\lambda = \mu_i^2$ para algún i. Por consiguiente, $\mu_i = \sqrt{\lambda}$ o $\mu_i = -\sqrt{\lambda}$; esto es, $\sqrt{\lambda}$ o $-\sqrt{\lambda}$ es un valor propio de A.

11.3. INVARIANCIA

Sea $T: V \to V$ lineal. Un subespacio W de V se dice *invariante bajo* T o *T-invariante* si T aplica W en sí mismo, o sea, si $v \in W$ implica $T(v) \in W$. En tal caso, T induce un operador lineal $\hat{T}: W \to W$ definido por $\hat{T}(w) = T(w)$ para todo $w \in W$.

EJEMPLO 11.2

a) Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal que gira cada vector alrededor del eje z un ángulo θ (Fig. 11-1); esto es, el definido por

$$T(x, y, z) = (x \cos \theta - y \operatorname{sen} \theta, x \operatorname{sen} \theta + y \cos \theta, z)$$

Obsérvese que cada vector $w = (a, b, 0)$ del plano xy, W, permanece en W bajo la aplicación T, luego W es T-invariante. Obsérvese asimismo que el eje z, U, es invariante bajo T. Además, la restricción de T a W gira cada vector alrededor del origen O, y la de T a U es la aplicación identidad de U.

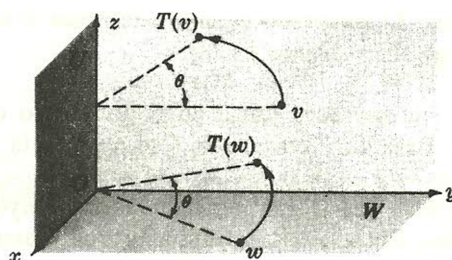


Figura 11-1.

b) Los vectores propios no nulos de un operador lineal $T: V \to V$ pueden caracterizarse como generadores de subespacios 1-dimensionales T-invariantes. Supongamos $T(v) = \lambda v$, $v \neq 0$. Entonces, $W = \{kv, k \in K\}$, el subespacio 1-dimensional generado por v, es invariante bajo T porque

$$T(kv) = kT(v) = k(\lambda v) = k\lambda v \in W$$

Recíprocamente, supongamos que $\dim U = 1$, que $u \neq 0$ genera U y que U es invariante bajo T. En ese caso, $T(u) \in U$ y así $T(u)$ es un múltiplo de u, o sea, $T(u) = \mu u$. De aquí que u sea un vector propio de T.

El próximo teorema, demostrado en el Problema 11.3, nos da una clase importante de subespacios invariantes.

Teorema 11.2: Sean $T: V \to V$ un operador lineal y $f(t)$ cualquier polinomio. El núcleo de $f(T)$ es invariante bajo T.

La noción de invariancia está relacionada con las representaciones matriciales (Problema 11.5) como sigue.

Teorema 11.3: Supongamos que W es un subespacio invariante de $T: V \to V$. Entonces T tiene una representación matricial por bloques $\begin{pmatrix} A & C \\ 0 & C \end{pmatrix}$, donde A es una representación matricial de la restricción $\hat{T}$ de T a W.

11.4. DESCOMPOSICIONES EN SUMA DIRECTA INVARIANTE

Un espacio vectorial V es la *suma directa* de sus subespacios $W_1, \ldots, W_r$, escrito

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

si todo vector $v \in V$ puede expresarse de forma única como

$$v = w_1 + w_2 + \cdots + w_r \qquad \text{con} \qquad w_i \in W_i$$

Es aplicable el siguiente teorema, demostrado en el Problema 11.7.

Teorema 11.4: Supongamos que $W_1, \ldots, W_r$ son subespacios de V y que

$$\{w_{11}, \ldots, w_{1n_1}\}, \ldots, \{w_{r1}, \ldots, w_{rn_r}\}$$

son bases de $W_1, \ldots, W_r$, respectivamente. En ese caso, V es la suma directa de los W_i si y sólo si la unión $B = \{w_{11}, \ldots, w_{1n_1}, \ldots, w_{r1}, \ldots, w_{rn_r}\}$ es una base de V.

Supongamos ahora que $T: V \to V$ es lineal y que V es la suma directa de subespacios T-invariantes (no nulos) $W_1, \ldots, W_r$:

$$V = W_1 \oplus \cdots \oplus W_r \qquad \text{y} \qquad T(W_i) \subseteq W_i \qquad i = 1, \ldots, r$$

Denotemos por T_i la restricción de T a W_i. Entonces se dice que T es *descomponible* en los operadores T_i o que T es la *suma directa* de los T_i, escrito $T = T_1 \oplus \cdots \oplus T_r$. Asimismo, se dice que los subespacios $W_1, \ldots, W_r$ *reducen* a T o que forman una *descomposición en suma directa T-invariante* de V.

Consideremos la situación especial en la que dos subespacios U y W reducen a un operador $T: V \to V$; sea, por ejemplo, $\dim U = 2$ y $\dim W = 3$ y supongamos que $\{u_1, u_2\}$ y $\{w_1, w_2, w_3\}$ son bases de U y W, respectivamente. Si T_1 y T_2 denotan las restricciones de T a U y W, tendremos

$$\begin{aligned} T_1(u_1) &= a_{11}u_1 + a_{12}u_2 \\ T_1(u_2) &= a_{21}u_1 + a_{22}u_2 \end{aligned} \qquad \begin{aligned} T_2(w_1) &= b_{11}w_1 + b_{12}w_2 + b_{13}w_3 \\ T_2(w_2) &= b_{21}w_1 + b_{22}w_2 + b_{23}w_3 \\ T_2(w_3) &= b_{31}w_1 + b_{32}w_2 + b_{33}w_3 \end{aligned}$$

De aquí

$$A = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

son representaciones matriciales de T_1 y T_2, respectivamente. Según el Teorema 11.4,

$$\{u_1, u_2, w_1, w_2, w_3\}$$

es una base de V. Dado que $T(u_i) = T_1(u_i)$ y $T(w_j) = T_2(w_j)$, la matriz de T en esta base es la matriz diagonal por bloques $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$.

Una generalización del argumento anterior proporciona el teorema enunciado a continuación.

Teorema 11.5: Supongamos que $T: V \to V$ es lineal y que V es la suma directa de subespacios T-invariantes, digamos $W_1, \ldots, W_r$. Si A_i es una representación matricial de la restricción de T a W_i, T puede representarse por la matriz diagonal por bloques

$$M = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & A_r \end{pmatrix}$$

La matriz diagonal por bloques M con entradas diagonales $A_1, \ldots, A_r$ se denomina a veces la suma directa de las matrices $A_1, \ldots, A_r$ y se denota por $M = A_1 \oplus \cdots \oplus A_r$.

11.5. DESCOMPOSICION PRIMARIA

El siguiente teorema muestra que un operador $T: V \to V$ es descomponible en operadores cuyos polinomios mínimos son potencias de polinomios irreducibles. Este es el primer paso hacia la obtención de una forma canónica de T.

Teorema de descomposición primaria 11.6: Sea $T: V \to V$ un operador lineal con polinomio mínimo

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_r(t)^{n_r}$$

donde los $f_i(t)$ son polinomios irreducibles normalizados distintos. Entonces V es la suma directa de subespacios T-invariantes $W_1, \ldots, W_r$, donde W_i es el núcleo de $f_i(T)^{n_i}$. Más aún, $f_i(t)^{n_i}$ es el polinomio mínimo de la restricción de T a W_i.

Siendo los polinomios $f_i(t)^{n_i}$ primos entre sí, el resultado fundamental precedente deriva (Problema 11.11) de los próximos dos teoremas.

Teorema 11.7: Supongamos que $T: V \to V$ es lineal y que $f(t) = g(t)h(t)$ son polinomios tales que $f(T) = \mathbf{0}$ y $g(t)$ y $h(t)$ son primos entre sí. En ese caso, V es la suma directa de los subespacios T-invariantes U y W, con $U = \operatorname{Ker} g(T)$ y $W = \operatorname{Ker} h(T)$.

Teorema 11.8: En el Teorema 11.7, si $f(t)$ es el polinomio mínimo de T [y $g(t)$ y $h(t)$ están normalizados], $g(t)$ y $h(t)$ son los polinomios mínimos de las restricciones de T *a* U y W, respectivamente.

Utilizaremos también el teorema de descomposición primaria para demostrar la útil caracterización de los operadores diagonalizables que ahora escribimos (véase la demostración en el Problema 11.12).

Teorema 11.9: Un operador lineal $T: V \to V$ es diagonalizable si y sólo si su polinomio mínimo es un producto de polinomios lineales diferentes.

Teorema 11.9 (forma alternativa): Una matriz A es similar a una matriz diagonal si y sólo si su polinomio mínimo es un producto de polinomios lineales diferentes.

EJEMPLO 11.3. Supongamos que $A \neq I$ es una matriz cuadrada para la cual $A^3 = I$. Determinemos si A es o no similar a una matriz diagonal, cuando A es una matriz sobre i) el cuerpo real $\mathbf{R}_i$, ii) el cuerpo complejo $\mathbf{C}$.

Puesto que $A^3 = I$, A es un cero del polinomio $f(t) = t^3 - 1 = (t-1)(t^2 + t + 1)$. El polinomio mínimo $m(t)$ de A no puede ser $t - 1$, ya que $A \neq I$. Por tanto,

$$m(t) = t^2 + t + 1 \qquad \text{o} \qquad m(t) = t^2 - 1$$

Dado que ninguno de los polinomios es producto de polinomios lineales sobre $\mathbf{R}$, A no es diagonalizable sobre $\mathbf{R}$. Por el contrario, cada uno de los polinomios es producto de polinomios lineales distintos sobre $\mathbf{C}$, luego A es diagonalizable sobre $\mathbf{C}$.

11.6. OPERADORES NILPOTENTES

Se dice que un operador lineal $T: V \to V$ es *nilpotente* si $T^n = 0$ para algún entero positivo n; llamamos a k el *índice de nilpotencia* de T si $T^k = 0$ pero $T^{k-1} \neq 0$. Análogamente, una matriz cuadrada A se dice nilpotente si $A^n = 0$ para algún entero positivo n, y de índice k si $A^k = 0$ pero $A^{k-1} \neq 0$. Claramente, el polinomio mínimo de un operador (matriz) nilpotente de índice k es $m(t) = t^k$, por lo que 0 es su único valor propio.

El resultado fundamental sobre operadores nilpotentes es el siguiente.

Teorema 11.10: Sea $T: V \to V$ un operador nilpotente de índice k. En tal caso, T tiene una representación matricial diagonal por bloques cuyas entradas diagonales son de la forma

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

(es decir, todas las entradas de N son 0 excepto aquellas situadas justo sobre la diagonal principal, que son 1). Hay al menos una N de orden k, siendo todas las demás de orden $\leq k$. El número de las N de cada orden posible está unívocamente determinado por T. Más aún, el número total de las N de todos los órdenes coincide con la nulidad de T.

En la demostración del teorema anterior (Problema 11.16) probaremos que el número de las N de orden i es igual a $2m_i - m_{i+1} - m_{i-1}$, donde m_i es la nulidad de T^i.

Señalamos que la matriz N precedente es a su vez nilpotente y que su índice de nilpotencia es igual a su orden (Problema 11.13). Nótese que la matriz N de orden 1 es justamente la matriz cero (0) 1×1.

11.7. FORMA CANONICA DE JORDAN

Un operador T puede ponerse en forma canónica de Jordan si sus polinomios característico y mínimo se factorizan en polinomios lineales. Esto siempre se cumple si K es el cuerpo complejo **C**. En cualquier caso, siempre podemos extender el cuerpo base K a un cuerpo en el que los polinomios característico y mínimo se factoricen en polinomios lineales; de modo que, en un sentido amplio, todo operador tiene una forma canónica de Jordan. Análogamente, toda matriz A es similar a una matriz en forma canónica de Jordan.

Teorema 11.11: Sea $T: V \to V$ un operador lineal cuyos polinomios mínimo y característico son, respectivamente,

$$\Delta(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \qquad \text{y} \qquad m(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}$$

donde los λ_i son escalares distintos. Entonces T tiene una representación matricial diagonal por bloques J, con entradas diagonales

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}$$

Por cada λ_i, los bloques correspondientes J_{ij} tienen las siguientes propiedades:

i) Existe al menos un J_{ij} de orden m_i; todos los demás J_{ij} son de orden $\leq m_i$.

ii) La suma de los órdenes de los J_{ij} es n_i.

iii) El número de los J_{ij} coincide con la multiplicidad geométrica de λ_i.

iv) El número de los J_{ij} de cada orden posible está unívocamente determinado por T.

La matriz J del teorema anterior recibe el nombre de *forma canónica de Jordan* del operador T. Cada bloque diagonal J_{ij} se llama *bloque de Jordan* perteneciente al valor propio λ_i. Obsérvese que

$$\begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Esto es,

$$J_{ij} = \lambda_i I + N$$

donde N es el bloque nilpotente que aparecía en el Teorema 11.10. De hecho, demostraremos el Teorema 11.11 (Problema 11.18) probando que T puede descomponerse en operadores suma, cada uno de ellos, de un operador escalar y un operador nilpotente.

EJEMPLO 11.4. Supongamos que los polinomios característico y mínimo de un operador T son, respectivamente,

$$\Delta(t) = (t-2)^4(t-3)^3 \qquad \text{y} \qquad m(t) = (t-2)^2(t-3)^2$$

La forma canónica de Jordan de T es, pues, una de las matrices:

$$\begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & 1 & & & \\ & & 0 & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & & & & \\ & & & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix}$$

Será la primera si T tiene dos vectores propios independientes pertenecientes a su valor propio 2; y será la segunda si T tiene tres vectores propios independientes correspondientes a 2.

11.8. SUBESPACIOS CICLICOS

Sea T un operador lineal en un espacio V de dimensión finita sobre K. Supongamos $v \in V$ y $v \neq 0$. El conjunto de todos los vectores de la forma $f(T)(v)$, donde $f(t)$ recorre todos los polinomios sobre K, es un subespacio T-invariante de V denominado el *subespacio T-cíclico* de *V generado por v*. Lo denotaremos por $Z(v, T)$, y denotaremos la restricción de T a $Z(v, T)$ por T_v. Podríamos definir equivalentemente $Z(v, T)$ como la intersección de todos los subespacios T-invariantes de V que contengan v.

Consideremos ahora la sucesión

$$v,\ T(v),\ T^2(v),\ T^3(v),\ \ldots$$

de potencias de T actuando sobre v. Sea k el menor entero tal que $T^k(v)$ es una combinación lineal de los vectores que lo preceden en la sucesión, digamos

$$T^k(v) = -a_{k-1}T^{k-1}(v) - \cdots - a_1 T(v) - a_0 v$$

En tal caso,

$$m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0$$

es el único polinomio normalizado de grado mínimo para el que $m_v(T)(v) = 0$. Llamamos a $m_v(t)$ el *T-aniquilador* de v y $Z(v, T)$.

Es aplicable el siguiente teorema (demostrado en el Problema 11.29).

Teorema 11.12: Definamos $Z(v, T)$, T_v y $m_v(t)$ como antes. Se cumple:

i) El conjunto $\{v, T(v), \ldots, T^{k-1}(v)\}$ es una base de $Z(v, T)$, luego $\dim Z(v, T) = k$.

ii) El polinomio mínimo de T_v es $m_v(t)$.

iii) La representación matricial de T_v en la base anterior es justamente la matriz acompañante (Ejemplo 8.12) de $m_v(t)$:

$$C = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \ldots & 1 & -a_{k-1} \end{pmatrix}$$

11.9. FORMA CANONICA RACIONAL

En esta sección presentamos la forma canónica racional para un operador lineal $T: V \to V$. Subrayamos que esta forma existe incluso cuando el polinomio mínimo no puede factorizarse en polinomios lineales. (Recordemos que no es este el caso para la forma canónica de Jordan.)

Lema 11.13: Sea $T: V \to V$ un operador lineal cuyo polinomio mínimo es $f(t)^n$, donde $f(t)$ es un polinomio irreducible normalizado. Entonces V es la suma directa.

$$V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$$

de subespacio T-cíclicos $Z(v_i, T)$ con correspondientes T-aniquiladores

$$f(t)^{n_1}, f(t)^{n_2}, \ldots, f(t)^{n_r} \qquad n = n_1 \geq n_2 \geq \cdots \geq n_r$$

Cualquier otra descomposión de V en subespacios T-cíclicos tiene el mismo número de componentes y el mismo conjunto de T-aniquiladores.

Señalamos que el lema precedente, demostrado en el Problema 11.31, no dice que los vectores v_i o los subespacios T-cíclicos $Z(v_i, T)$ estén unívocamente determinados por T, sino que el conjunto de T-aniquiladores lo está. Así T tiene una única representación matricial

$$\begin{pmatrix} C_1 & & & \\ & C_2 & & \\ & & \cdots & \\ & & & C_r \end{pmatrix}$$

en la que las C_i son matrices acompañantes. De hecho, las C_i son las matrices acompañantes de los polinomios $f(t)^{n_i}$.

Utilizando el teorema de descomposición primaria y el Lema 11.13 obtenemos el resultado fundamental que sigue.

Teorema 11.14: Sea $T: V \to V$ un operador lineal con polinomio mínimo

$$m(t) = f_1(t)^{m_1} f_2(t)^{m_2} \cdots f_s(t)^{m_s}$$

donde los $f_i(t)$ son polinomios irreducibles normalizados distintos. En ese caso, T tiene una única representación matricial diagonal por bloques

$$\begin{pmatrix} C_{11} & & & & & & \\ & \cdots & & & & & \\ & & C_{1r_1} & & & & \\ & & & \cdots & & & \\ & & & & C_{s1} & & \\ & & & & & \cdots & \\ & & & & & & C_{sr_s} \end{pmatrix}$$

en la que las C_{ij} son matrices acompañantes. En concreto, las C_{ij} son las matrices acompañantes de los polinomios $f_i(t)^{n_{ij}}$, con

$$m_1 = n_{11} \geq n_{12} \geq \cdots \geq n_{1r_1}, \ldots, m_s = n_{s1} \geq n_{s2} \geq \cdots \geq n_{sr_s}$$

La anterior representación matricial de T se llama su *forma canónica racional*. Los polinomios $f_i(t)^{n_{ij}}$ se conocen como *divisores elementales* de T.

EJEMPLO 11.5. Sea V un espacio vectorial de dimensión 6 sobre $\mathbf{R}$ y T un operador lineal cuyo polinomio mínimo es $m(t) = (t^2 - t + 3)(t - 2)^2$. La forma canónica racional de T es una de las sumas directas de matrices acompañantes:

i) $C(t^2 - t + 3) \oplus C(t^2 - t + 3) \oplus C((t - 2)^2)$

ii) $C(t^2 - t + 3) \oplus C((t - 2)^2) \oplus C((t - 2)^2)$

iii) $C(t^2 - t + 3) \oplus C((t - 2)^2) \oplus C(t - 2) \oplus C(t - 2)$

donde $C(f(t))$ es la matriz acompañante de $f(t)$; esto es,

$$\text{i)}\quad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -3 & & \\ & & 1 & 1 & & \\ & & & & 0 & -4 \\ & & & & 1 & 4 \end{pmatrix} \qquad \text{ii)}\quad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -4 & & \\ & & 1 & 4 & & \\ & & & & 0 & -4 \\ & & & & 1 & 4 \end{pmatrix} \qquad \text{iii)}\quad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -4 & & \\ & & 1 & 4 & & \\ & & & & 2 & \\ & & & & & 2 \end{pmatrix}$$

11.10. ESPACIOS COCIENTE

Sean V un espacio vectorial sobre un cuerpo K y W un subespacio de V. *Si* v es cualquier vector en V, escribimos $v + W$ para representar el conjunto de sumas $v + w$ con $w \in W$:

$$v + W = \{v + w : w \in W\}$$

Estos conjuntos se denominan *variedades afines* de W en V. Probaremos (Problema 11.22) que dichas variedades afines constituyen una partición de V en subconjuntos mutuamente disjuntos.

EJEMPLO 11.6. Sea W el subespacio de $\mathbf{R}^2$ definido como

$$W = \{(a, b) : a = b\}$$

Es decir, W es la recta dada por la ecuación $x - y = 0$. Podemos ver $v + W$ como una traslación de la recta, obtenida sumando v a cada punto de W. Como se muestra en la Figura 11-2, $v + W$ es también una recta y es paralela a W. De este modo, las variables afines de W en $\mathbf{R}^2$ son precisamente todas las rectas paralelas a W.

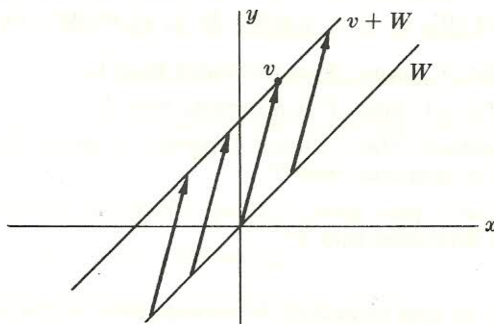


Figura 11-2.

En el teorema enunciado a continuación hacemos uso de las variedades afines de un subespacio W de un espacio vectorial V para definir un nuevo espacio vectorial; se llama el *espacio cociente* de V por W y se denota por V/W.

Teorema 11.15: Sea W un subespacio de un espacio vectorial V sobre un cuerpo K. Las variedades afines de W en V forman un espacio vectorial sobre K con las siguientes operaciones de suma y producto por un escalar:

i) $(u + W) + (v + W) = (u + v) + W$.

ii) $k(u + W) = ku + W$, donde $k \in K$.

Hacemos notar que en la demostración del Teorema 11.15 (Problema 11.24) es necesario probar primero que las operaciones están bien definidas; o sea, que siempre que $u + W = u' + W$ y $v + W = v' + W$ se cumple:

i) $(u + v) + W = (u' + v') + W$ y ii) $ku + W = ku' + W$ para cualquier $k \in K$

En el caso de un subespacio invariante, disponemos del útil resultado, demostrado en el Problema 11.27, que enseguida escribimos.

Teorema 11.16: Supongamos que W es un subespacio invariante bajo un operador lineal $T: V \to V$. Entonces T induce un operador lineal $\bar{T}$ en V/W definido por $\bar{T}(v + W) = T(v) + W$. Más aún, si T es un cero de algún polinomio, también lo es $\bar{T}$. El polinomio mínimo de $\bar{T}$ es divisor, pues, del de T.

PROBLEMAS RESUELTOS

SUBESPACIOS INVARIANTES

11.1. Supóngase que $T: V \to V$ es lineal. Probar que cada uno de los siguientes subespacios es invariante bajo T: *a*) $\{0\}$, *b*) V, *c*) núcleo de T, *d*) imagen de T.

a) Tenemos $T(0) = 0 \in \{0\}$, luego $\{0\}$ es invariante bajo T.

b) Para todo $v \in V$, $T(v) \in V$, luego V es invariante bajo T.

c) Sea $u \in \operatorname{Ker} T$. Entonces $T(u) = 0 \in \operatorname{Ker} T$ porque el núcleo de T es un subespacio de V. Siendo así, Ker T es invariante bajo T.

d) Dado que $T(v) \in \operatorname{Im} T$ para todo $v \in V$, es claramente cierto cuando $v \in \operatorname{Im} T$. Por tanto, la imagen de T es invariante bajo T.

11.2. Supóngase que $\{W_i\}$ es una colección de subespacios T-invariantes de un espacio vectorial V. Probar que la intersección $W = \cap_i W_i$ es también T-invariante.

Supongamos $v \in W$; en tal caso, $v \in W_i$ para todo i. Puesto que W_i es T-invariante, $T(v) \in W_i$ para todo i. De este modo, $T(v) \in W = \cap_i W_i$ y W es, pues, T-invariante

11.3. Demostrar el Teorema 11.2.

Supongamos $v \in \operatorname{Ker} f(T)$, es decir, $f(T)(v) = 0$. Debemos probar que $T(v)$ también pertenece al núcleo de $f(T)$, o sea, $f(T)(T(v)) = 0$. Dado que $f(t)t = tf(t)$, tenemos $f(T)T = Tf(T)$. Así pues,

$$f(T)T(v) = Tf(T)(v) = T(0) = 0$$

como se pedía.

11.4. Hallar todos los subespacios invariantes de $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$ vista como un operador en $\mathbf{R}^2$.

Por el Problema 11.1, $\mathbf{R}^2$ y $\{0\}$ son invariantes bajo A. Ahora bien, si A tiene cualquier otro subespacio invariante, éste debe ser 1-dimensional. Sin embargo, el polinomio característico de A es

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 5 \\ -1 & t+2 \end{vmatrix} = t^2 + 1$$

Por esta razón, A no tiene valores propios (en $\mathbf{R}$) ni por tanto vectores propios. Pero los subespacios invariantes 1-dimensionales corresponden a los vectores propios; así $\mathbf{R}^2$ y $\{0\}$ son los únicos subespacios invariantes bajo A.

11.5. Demostrar el Teorema 11.3.

Escogemos una base $\{w_1, \ldots, w_r\}$ de W y la extendemos a una $\{w_1, \ldots, w_r, v_1, \ldots, v_s\}$ de V. Tenemos

$$\begin{aligned}
\hat{T}(w_1) = T(w_1) &= a_{11}w_1 + \cdots + a_{1r}w_r \\
\hat{T}(w_2) = T(w_2) &= a_{21}w_1 + \cdots + a_{2r}w_r \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\hat{T}(w_r) = T(w_r) &= a_{r1}w_1 + \cdots + a_{rr}w_r \\
T(v_1) &= b_{11}w_1 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s \\
T(v_2) &= b_{21}w_1 + \cdots + b_{2r}w_r + c_{21}v_1 + \cdots + c_{2s}v_s \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(v_s) &= b_{s1}w_1 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}v_s
\end{aligned}$$

Pero la matriz de T en esta base es la traspuesta de la matriz de coeficientes del sistema de ecuaciones anterior. (Véase la página 407.) Por consiguiente, es de la forma $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, donde A es la traspuesta de la matriz de coeficientes del subsistema obvio. Por el mismo argumento, A es la matriz de $\hat{T}$ relativa a la base $\{w_i\}$ de W.

11.6. Denotemos por $\hat{T}$ la restricción de un operador T a un subespacio invariante W, es decir, $\hat{T}(w) = T(w)$ para todo $w \in W$. Demostrar:

i) Para cualquier polinomio $f(t)$, $f(\hat{T})(w) = f(T)(w)$.

ii) El polinomio mínimo de $\hat{T}$ es divisor del polinomio mínimo de T.

i) Si $f(t)$ es de grado 0 o de grado 1, el resultado claramente se cumple.

Supongamos que grado $f = n > 1$ y que el resultado se cumple para polinomios de grado menor que n. Sea

$$f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

En ese caso,

$$\begin{aligned} f(\hat{T})(w) &= (a_n \hat{T}^n + a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) = \\ &= (a_n \hat{T}^{n-1})(\hat{T}(w)) + (a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) = \\ &= (a_n T^{n-1})(T(w)) + (a_{n-1}T^{n-1} + \cdots + a_0 I)(w) = \\ &= f(T)(w) \end{aligned}$$

ii) Sea $m(t)$ el polinomio mínimo de T. Por i), $m(\hat{T})(w) = m(T)(w) = \mathbf{0}(w) = 0$ para todo $w \in W$; esto es, $\hat{T}$ es un cero del polinomio $m(t)$. De aquí que el polinomio mínimo de $\hat{T}$ sea divisor de $m(t)$.

DESCOMPOSICIONES EN SUMA DIRECTA INVARIANTE

11.7. Demostrar el Teorema 11.4.

Supongamos que B es una base de V. Entonces, para todo $v \in V$,

$$v = a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} + \cdots + a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = w_1 + w_2 + \cdots + w_r$$

donde $w_i = a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} \in W_i$. A continuación demostramos que tal suma es única. Supongamos

$$v = w'_1 + w'_2 + \cdots + w'_r \qquad \text{con } w'_i \in W_i$$

Como $\{w_{i1}, \ldots, w_{in_i}\}$ es una base de W_i, $w'_i = b_{i1}w_{i1} + \cdots + b_{in_i}w_{in_i}$ y por tanto

$$v = b_{11}w_{11} + \cdots + b_{1n_1}w_{1n_1} + \cdots + b_{r1}w_{r1} + \cdots + b_{rn_r}w_{rn_r}$$

Al ser B una base de V, $a_{ij} = b_{ij}$ para cada i y cada j. Por esta razón, $w_i = w'_i$ y la suma para v es, pues, única. En consecuencia, V es la suma directa de los W_i.

Recíprocamente, supongamos que V es la suma directa de los W_i. En ese caso, para todo $v \in V$, $v = w_1 + \cdots + w_r$, donde $w_i \in W_i$. Siendo $\{w_{ij}\}$ una base de W_i, cada w_i es combinación lineal de los w_{ij}, de modo que v es combinación lineal de los elementos de B. Probemos ahora que B es linealmente independiente. Supongamos

$$a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} + \cdots + a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = 0$$

Nótese que $a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} \in W_i$. Tenemos, asimismo, que $0 = 0 + 0 + \cdots + 0$ con $0 \in W_i$. Dado que tal suma para 0 es única,

$$a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} = 0 \qquad \text{para } i = 1, \ldots, r$$

La independencia de las bases $\{w_{ij}\}$ implica que todas las a son 0. Así B es linealmente independiente y por ende una base de V.

11.8. Supóngase que $T: V \to V$ es lineal y que $T = T_1 \oplus T_2$ con respecto a una descomposición en suma directa T-invariante $V = U \oplus W$. Mostrar que:

a) $m(t)$ es el mínimo común múltiplo de $m_1(t)$ y $m_2(t)$, donde $m(t)$, $m_1(t)$ y $m_2(t)$ son los polinomios mínimos de T, T_1 y T_2, respectivamente.

b) $\Delta(t) = \Delta_1(t)\Delta_2(t)$, donde $\Delta(t)$, $\Delta_1(t)$ y $\Delta_2(t)$ son los polinomios característicos de T, T_1 y T_2, respectivamente.

a) De acuerdo con el Problema 11.6, $m_1(t)$ y $m_2(t)$ son divisores de $m(t)$. Supongamos ahora que $f(t)$ es múltiplo de $m_1(t)$ y $m_2(t)$ simultáneamente; en tal caso, $f(T_1)(U) = 0$ y $f(T_2)(W) = 0$. Sea $v \in V$; entonces $v = u + w$ con $u \in U$ y $w \in W$. Ahora,

$$f(T)v = f(T)u + f(T)w = f(T_1)u + f(T_2)w = 0 + 0 = 0$$

Es decir, T es un cero de $f(t)$. Por ello, $m(t)$ es divisor de $f(t)$ y así $m(t)$ es el mínimo común múltiplo de $m_1(t)$ y $m_2(t)$.

b) En virtud del Teorema 11.5, T tiene una representación matricial $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, siendo A y B las representaciones matriciales de T_1 y T_2, respectivamente. Entonces

$$\Delta(t) = |tI - M| = \begin{vmatrix} tI - A & 0 \\ 0 & tI - B \end{vmatrix} = |tI - A||tI - B| = \Delta_1(t)\,\Delta_2(t)$$

como se pedía.

11.9. Demostrar el Teorema 11.7.

Nótese primero que U y W son T-invariantes, según el Teorema 11.2. Al ser $g(t)$ y $h(t)$ primos entre sí, existen polinomios $r(t)$ y $s(t)$ tales que

$$r(t)g(t) + s(t)h(t) = 1$$

Por consiguiente, para el operador T,

$$r(T)g(T) + s(T)h(T) = 1 \qquad (*)$$

Sea $v \in V$; por $(*)$,

$$v = r(T)g(T)v + s(T)h(T)v$$

Pero el primer término en esta suma pertenece a $W = \text{Ker}\, h(T)$. De hecho, puesto que la composición de polinomios en T es conmutativa,

$$h(T)r(T)g(T)v = r(T)g(T)h(T)v = r(T)f(T)v = r(T)0v = 0$$

Similarmente, el segundo término pertenece a U. Por tanto, V es la suma de U y W.

Para demostrar que $V = U \oplus W$ debemos probar que la suma $v = u + w$ con $u \in U$, $w \in W$, está unívocamente determinada por v. Aplicando el operador $r(T)g(T)$ a $v = u + w$ y utilizando $g(T)u = 0$ obtenemos

$$r(T)g(T)v = r(T)g(T)u + r(T)g(T)w = r(T)g(T)w$$

Asimismo, aplicando $(*)$ a w sólo y usando $h(T)w = 0$,

$$w = r(T)g(T)w + s(T)h(T)w = r(T)g(T)w$$

Las dos últimas ecuaciones dan, conjuntamente, $w = r(T)g(T)v$, de modo que w está unívocamente determinado por v. Análogamente, u está unívocamente determinado por v. De aquí $V = U \oplus W$, como se pedía.

11.10. Demostrar el Teorema 11.8: En el Teorema 11.7 (Problema 11.9), si $f(t)$ es el polinomio mínimo de T [y $g(t)$ y $h(t)$ están normalizados], $g(t)$ es el polinomio mínimo de la restricción $\hat{T}_1$ de T a U y $h(t)$ el de la restricción $\hat{T}_2$ de T a W.

Sean $m_1(t)$ y $m_2(t)$ los polinomios mínimos de $\hat{T}_1$ y $\hat{T}_2$, respectivamente. Nótese que $g(\hat{T}_1) = \mathbf{0}$ y $h(\hat{T}_2) = \mathbf{0}$ porque $U = \text{Ker}\, g(T)$ y $W = \text{Ker}\, h(T)$. De este modo,

$$m_1(t) \text{ divide } g(t) \qquad \text{y} \qquad m_2(t) \text{ divide } h(t) \qquad [1]$$

Según el Problema 11.9, $f(t)$ es el mínimo común múltiplo de $m_1(t)$ y $m_2(t)$. Pero $m_1(t)$ y $m_2(t)$ son primos ente sí, por serlo $g(t)$ y $h(t)$. En consecuencia, $f(t) = m_1(t)m_2(t)$. Tenemos también que $f(t) = g(t)h(t)$. Estas dos ecuaciones, junto con [1] y el hecho de que todos los polinomios están normalizados, implican que $g(t) = m_1(t)$ y $h(t) = m_2(t)$, como se deseaba probar.

11.11. Demostrar el Teorema de Descomposición Primaria 11.6.

Se demuestra por inducción en r. El caso $r = 1$ es trivial. Supongamos que se ha demostrado el teorema para $r - 1$. De acuerdo con el Teorema 11.7, podemos escribir V como la suma directa de subespacios T-invariantes W_1 y V_1, donde W_1 es el núcleo de $f_1(T)^{n_1}$ y V_1 el de $f_2(T)^{n_2} \cdots f_r(T)^{n_r}$. Por el Teorema 11.8, los polinomios mínimos de las restricciones de T a W_1 y V_1 son, respectivamente, $f_1(t)^{n_1}$ y $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$.

Denotemos la restricción de T a V_1 por $\hat{T}_1$. Por la hipótesis de inducción, V_1 es la suma directa de subespacios $W_2, \ldots, W_r$ tales que W_i es el núcleo de $f_i(\hat{T}_1)^{n_i}$ y $f_i(t)^{n_i}$ el polinomio mínimo de la restricción de $\hat{T}_1$ a W_i. Pero el núcleo de $f_i(T)^{n_i}$ para $i = 2, \ldots, r$ está necesariamente contenido en V_1 porque $f_i(t)^{n_i}$ es divisor de $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$. El núcleo de $f_i(T)^{n_i}$ coincide, pues, con el de $f_i(\hat{T}_1)^{n_i}$, que es W_i. Asimismo, la restricción de T a W_i es la misma que la de $\hat{T}_1$ a W_i (para $i = 2, \ldots, r$), luego $f_i(t)^{n_i}$ es también el polinomio mínimo de la restricción de T a W_i. De esta manera, $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ es la descomposición de T deseada.

11.12. Demostrar el Teorema 11.9.

Supongamos que $m(t)$ es un producto de polinomios lineales distintos; por ejemplo,

$$m(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_r)$$

donde los λ_i son escalares diferentes. En virtud del Teorema 11.6, V es la suma directa de subespacios $W_1, \ldots, W_r$, con $W_i = \text{Ker}\,(T - \lambda_i I)$. Así pues, si $v \in W_i$, entonces $(T - \lambda_i I)v = 0$ o $T(v) = \lambda_i v$. En otras palabras, todo vector en W_i es un vector propio perteneciente al valor propio λ_i. Según el Teorema 11.4, la unión de bases de $W_1, \ldots, W_r$ es una base de V. Esta base está constituida por vectores propios, por lo que T es diagonalizable.

Recíprocamente, supongamos que T es diagonalizable, es decir, que V tiene una base consistente en vectores propios de T. Sean $\lambda_1, \ldots, \lambda_s$ los valores propios distintos de T. El operador

$$f(T) = (T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_s I)$$

aplica, pues, cada vector de la base en 0. De este modo, $f(T) = 0$ y por tanto el polinomio mínimo $m(t)$ de T es divisor del polinomio

$$f(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_s I)$$

En consecuencia, $m(t)$ es un producto de polinomios lineales distintos.

OPERADORES NILPOTENTES. FORMA CANONICA DE JORDAN

11.13. Sea $T\colon V \to V$ lineal. Supóngase, para $v \in V$, $T^k(v) = 0$ pero $T^{k-1}(v) \neq 0$. Demostrar:

a) El conjunto $S = \{v, T(v), \ldots, T^{k-1}(v)\}$ es linealmente independiente.

b) El subespacio W generado por S es T-invariante.

c) La restricción $\hat{T}$ de T a W es nilpotente de índice k.

d) Respecto a la base $\{T^{k-1}(v), \ldots, T(v), v\}$ de W, la matriz de T es de la forma

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

Por consiguiente, la matriz k-cuadrada anterior es nilpotente de índice k.

a) Supongamos

$$av + a_1 T(v) + a_2 T^2(v) + \cdots + a_{k-1} T^{k-1}(v) = 0 \qquad (*)$$

Aplicando T^{k-1} a $(*)$ y usando $T^k(v) = 0$ obtenemos $aT^{k-1}(v) = 0$; como $T^{k-1}(v) \neq 0$, $a = 0$. Ahora, aplicando T^{k-2} a $(*)$ y usando $T^k(v) = 0$ y $a = 0$, encontramos que $a_1 T^{k-1}(v) = 0$, luego $a_1 = 0$. Aplicando a continuación T^{k-3} a $(*)$ y usando $T^k(v) = 0$ y $a = a_1 = 0$ obtenemos $a_2 T^{k-1}(v) = 0$, luego $a_2 = 0$. Continuando este proceso encontramos que todos los a son 0; por tanto, S es independiente.

b) Sea $v \in W$. Entonces

$$v = bv + b_1 T(v) + b_2 T^2(v) + \cdots + b_{k-1} T^{k-1}(v)$$

Empleando $T^k(v) = 0$ tenemos que

$$T(v) = bT(v) + b_1 T^2(v) + \cdots + b_{k-2} T^{k-1}(v) \in W$$

Así W es T-invariante.

c) Por hipótesis, $T^k(v) = 0$. Por eso, para $i = 0, \ldots, k-1$,

$$\hat{T}^k(T^i(v)) = T^{k+i}(v) = 0$$

Esto es, aplicando $\hat{T}^k$ a cada generador de W, obtenemos 0; por este motivo, $\hat{T}^k = \mathbf{0}$, de manera que $\hat{T}$ es nilpotente de índice k a lo sumo. Por otra parte, $\hat{T}^{k-1}(v) = T^{k-1}(v) \neq 0$; de aquí que $\hat{T}$ sea nilpotente de índice k exactamente.

d) Para la base $\{T^{k-1}(v), T^{k-2}(v), \ldots, T(v), v\}$ de W,

$$\begin{aligned} \hat{T}(T^{k-1}(v)) &= T^k(v) = 0 \\ \hat{T}(T^{k-2}(v)) &= T^{k-1}(v) \\ \hat{T}(T^{k-3}(v)) &= T^{k-2}(v) \\ &\cdots\cdots\cdots\cdots \\ \hat{T}(T(v)) &= T^2(v) \\ \hat{T}(v) &= T(v) \end{aligned}$$

La matriz de T en esta base es, pues,

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

11.14. Sea $T: V \to V$ lineal. Sean $U = \text{Ker } T^i$ y $W = \text{Ker } T^{i+1}$. Mostrar que: *a*) $U \subseteq W$, *b*) $T(W) \subseteq U$.

a) Supongamos $u \in U = \text{Ker } T^i$. En tal caso, $T^i(u) = 0$, por lo que $T^{i+1}(u) = T(T^i(u)) = T(0) = 0$. De este modo, $u \in \text{Ker } T^{i+1} = W$. Pero esto es cierto para todo $u \in U$, luego $U \subseteq W$.

b) De forma similar, si $w \in W = \text{Ker } T^{i+1}$, $T^{i+1}(w) = 0$. Siendo así, $T^{i+1}(w) = T^i(T(w)) = T^i(0) = 0$, por lo que $T(W) \subseteq U$.

11.15. Sea $T: V \to V$ lineal. Sean $X = \text{Ker } T^{i-1}$, $Y = \text{Ker } T^{i-1}$ y $Z = \text{Ker } T^i$. Por el Problema 11.14, $X \subseteq Y \subseteq Z$. Supóngase que

$$\{u_1, \ldots, u_r\} \qquad \{u_1, \ldots, u_r, v_1, \ldots, v_s\} \qquad \{u_1, \ldots, u_r, v_1, \ldots, v_s\, w_1, \ldots, w_t\}$$

son bases de X, Y y Z, respectivamente. Mostrar que

$$S = \{u_1, \ldots, u_r, T(w_1), \ldots, T(w_t)\}$$

está contenido en Y y es linealmente independiente.

Por el Problema 11.4, $T(Z) \subseteq Y$ y por tanto $S \subseteq Y$. Supongamos ahora que S es linealmente dependiente. Existe entonces una relación

$$a_1u_1 + \cdots + a_ru_r + b_1T(w_1) + \cdots + b_tT(w_t) = 0$$

donde al menos uno de los coeficientes no es cero. Además, dado que $\{u_i\}$ es independiente, al menos uno de los b_k debe ser no nulo. Trasponiendo hallamos

$$b_1T(w_1) + \cdots + b_tT(w_t) = -a_1u_1 - \cdots - a_ru_r \in X = \text{Ker } T^{i-2}$$

Por consiguiente, $$T^{i-2}(b_1T(w_1) + \cdots + b_tT(w_t)) = 0$$

Así $$T^{i-1}(b_1w_1 + \cdots + b_tw_t) = 0 \qquad \text{luego} \qquad b_1w_1 + \cdots + b_tw_t \in Y = \text{Ker } T^{i-1}$$

Puesto que $\{u_i, v_j\}$ genera Y, obtenemos una relación entre los u_i, v_j y w_k en la que uno de los coeficientes, o sea, uno de los b_k, no es cero. Esto contradice el hecho de que $\{u_i, v_j, w_k\}$ es independiente. Por ello, S debe ser también independiente.

11.16. Demostrar el Teorema 11.10.

Supongamos $\dim V = n$. Sean $W_1 = \text{Ker } T$, $W_2 = \text{Ker } T^2, \ldots, W_k = \text{Ker } T^k$. Tomemos $m_i = \dim W_i$ para $i = 1, \ldots, k$. Por ser T de índice k, $W_k = V$ y $W_{k-1} \neq V$, de modo que $m_{k-1} < m_k = n$. Por el Problema 11.14,

$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_k = V$$

Así, por inducción, podemos elegir una base $\{u_1, \ldots, u_n\}$ de V tal que $\{u_1, \ldots, u_{m_i}\}$ es una base de W_i.

Escogemos ahora una nueva base de V con respecto a la cual T tenga la forma deseada. Resultará conveniente designar los miembros de esta base mediante pares de índices. Comenzamos tomando

$$v(1, k) = u_{m_{k-1}+1},\ v(2, k) = u_{m_{k-1}+2}, \ldots, v(m_k - m_{k-1}, k) = u_{m_k}$$

y

$$v(1, k-1) = Tv(1, k),\ v(2, k-1) = Tv(2, k), \ldots, v(m_k - m_{k-1}, k-1) = Tv(m_k - m_{k-1}, k)$$

Por el Problema 11.15,

$$S_1 = \{u_1, \ldots, u_{m_{k-2}}, v(1, k-1), \ldots, v(m_k - m_{k-1}, k-1)\}$$

es un subconjunto linealmente independiente de W_{k-1}. Extendemos S_1 a una base de W_{k-1} añadiendo nuevos elementos (si fuera necesario) que denotamos por

$$v(m_k - m_{k-1} + 1, k-1), v(m_k - m_{k-1} + 2, k-1), \ldots, v(m_{k-1} - m_{k-2}, k-1)$$

A continuación tomamos

$$v(1, k-2) = Tv(1, k-1), v(2, k-2) = Tv(2, k-1), \ldots, v(m_{k-1} - m_{k-2}, k-2) = Tv(m_{k-1} - m_{k-2}, k-1)$$

De nuevo por el Problema 11.15,

$$S_2 = \{u_1, \ldots, u_{m_{k-3}}, v(1, k-2), \ldots, v(m_{k-1} - m_{k-2}, k-2)\}$$

es un subconjunto linealmente independiente de W_{k-2} que podemos extender a una base de W_{k-2} añadiendo elementos

$$v(m_{k-1} - m_{k-2} + 1, k-2), v(m_{k-1} - m_{k-2} + 2, k-2), \ldots, v(m_{k-2} - m_{k-3}, k-2)$$

Continuando de esta manera conseguimos una nueva base de V que por conveniencia ordenamos como sigue:

$$\begin{array}{llll}
v(1, k), & \ldots, v(m_k - m_{k-1}, k) & & \\
v(1, k-1), & \ldots, v(m_k - m_{k-1}, k-1), & \ldots, v(m_{k-1} - m_{k-2}, k-1) & \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
v(1, 2), & \ldots, v(m_k - m_{k-1}, 2), & \ldots, v(m_{k-1} - m_{k-2}, 2), \ldots, v(m_2 - m_1, 2) & \\
v(1, 1), & \ldots, v(m_k - m_{k-1}, 1), & \ldots, v(m_{k-1} - m_{k-2}, 1), \ldots, v(m_2 - m_1, 1), \ldots, v(m_1, 1) &
\end{array}$$

La fila inferior forma una base de W_1, las dos inferiores forman una de W_2, etc. Pero lo que es realmente importante para nosotros es que T aplica cada vector en el situado inmediatamente bajo él en la tabla, o en 0 si el vector está en la última fila. Esto es,

$$Tv(i, j) = \begin{cases} v(i, j-1) & \text{para } j > 1 \\ 0 & \text{para } j = 1 \end{cases}$$

Ahora queda claro [véase el Problema 11.13 *a*)] que T tendrá la forma deseada si los $v(i, j)$ están ordenados lexicográficamente: empezando por $v(1, 1)$ y subiendo por la primera columna hasta $v(1, k)$, saltando después a $v(2, 1)$ y subiendo por la segunda columna tanto como sea posible, etc.

Más aún, habrá exactamente

$$\begin{array}{rl}
m_k - m_{k-1} & \text{entradas diagonales de orden } k \\
(m_{k-1} - m_{k-2}) - (m_k - m_{k-1}) = 2m_{k-1} - m_k - m_{k-2} & \text{entradas diagonales de orden } k-1 \\
\cdots\cdots & \cdots\cdots \\
2m_2 - m_1 - m_3 & \text{entradas diagonales de orden } 2 \\
2m_1 - m_2 & \text{entradas diagonales de orden } 1
\end{array}$$

como puede leerse directamente de la tabla. En particular, al estar los números $m_1, \ldots, m_k$ unívocamente determinados por T, el número de entradas diagonales de cada orden también lo está. Finalmente, la identidad

$$m_1 = (m_k - m_{k-1}) + (2m_{k-1} - m_k - m_{k-2}) + \cdots + (2m_2 - m_1 - m_3) + (2m_1 - m_2)$$

muestra que la nulidad m_1 de T es el número total de entradas diagonales de T.

11.17. Sea $A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Entonces $A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ y $A^3 = 0$; luego A es nilpotente de índice 3. Hallar la matriz nilpotente en forma canónica M que es similar a A.

Siendo A nilpotente de índice 3, M contendrá un bloque diagonal de orden 3 y ninguno de orden mayor. Nótese que rango $A = 2$; por tanto, nulidad $A = 5 - 2 = 3$. Así pues, M contiene tres bloques diagonales. En consecuencia, M debe contener un bloque diagonal de orden 3 y dos de orden 1; esto es,

$$M = \left(\begin{array}{ccc|c|c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

11.18. Demostrar el Teorema 11.11.

De acuerdo con el teorema de descomposición primaria, T es descomponible en operadores $T_1, \ldots, T_r$, es decir, $T = T_1 \oplus \cdots \oplus R_r$, donde $(t - \lambda_i)^{m_i}$ es el polinomio mínimo de T_i. De este modo, en particular,

$$(T_1 - \lambda_1 I)^{m_1} = \mathbf{0}, \ldots, (T_r - \lambda_r I)^{m_r} = \mathbf{0}$$

Tomemos $N_i = T_i - \lambda_i I$. Entonces, para $i = 1, \ldots, r$,

$$T_i = N_i + \lambda_i I \qquad \text{donde } N_i^{m_i} = \mathbf{0}$$

O sea, T_i es la suma del operador escalar $\lambda_i I$ y un operador nilpotente N_i, que es de índice m_i porque $(t - \lambda_i)^{m_i}$ es el polinomio mínimo de T_i.

Ahora, según el Teorema 11.10 referente a operadores nilpotentes, podemos escoger una base de manera que N_i esté en forma canónica. En esta base, $T_i = N_i + \lambda_i I$ se representa por una matriz diagonal por bloques M_i cuyas entradas diagonales son las matrices J_{ij}. La suma directa J de las matrices M_i está en forma canónica de Jordan y, por el Teorema 11.5, es una representación matricial de T.

Por último, debemos probar que los bloques J_{ij} satisfacen las propiedades requeridas. La propiedad i) deriva del hecho de que N_i es de índice m_i. La propiedad ii) es cierta, puesto que T y J tienen el mismo polinomio característico. La propiedad iii) es cierta, ya que la nulidad de $N_i = T_i - \lambda_i I$ es igual a la multiplicidad geométrica del valor propio λ_i. La propiedad iv) se desprende del hecho de que los T_i y por ende los N_i están unívocamente determinados por T.

11.19. Determinar todas las formas canónicas de Jordan posibles para un operador lineal $T: V \to V$ cuyo polinomio característico sea $\Delta(t) = (t-2)^3(t-5)^2$.

Dado que $t - 2$ tiene exponente 3 en $\Delta(t)$, 2 debe aparecer tres veces en la diagonal principal. De forma similar, 5 debe aparecer dos veces. Siendo así, las formas canónicas de Jordan posibles son

$$\begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix} \qquad \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix}$$

i) ii) iii)

$$\begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix} \qquad \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix}$$

iv) v) iv)

11.20. Determinar todas las formas canónicas de Jordan posibles para una matriz de orden 5 cuyo polinomio mínimo sea $m(t) = (t - 2)^2$.

J debe tener un bloque de Jordan de orden 2 y los otros deben ser de orden 2 ó 1. De este modo, hay solamente dos posibilidades:

$$J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & 1 & \\ & & & 2 & \\ & & & & 2 \end{pmatrix} \quad \text{o} \quad J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$$

Nótese que todas las entradas diagonales deben ser 2 porque 2 es el único valor propio.

ESPACIO COCIENTE Y FORMA TRIANGULAR

11.21. Sea W un subespacio de un espacio vectorial V. Probar que son equivalentes: i) $u \in v + W$, ii) $u - v \in W$, iii) $v \in u + W$.

Supongamos $u \in v + W$. Entonces existe $w_0 \in W$ tal que $u = v + w_0$. Por tanto, $u - v = w_0 \in W$. Recíprocamente, supongamos $u - v \in W$. En tal caso, $u - v = w_0$, donde $w_0 \in W$. Por tanto, $u = v + w_0 \in W$. Así pues, i) y ii) son equivalentes.

Asimismo tenemos: $u - v \in W$ si y sólo si $-(u - v) = v - u \in W$ si y sólo si $v \in u + W$. De esta manera, ii) y iii) son también equivalentes.

11.22. Demostrar la siguiente afirmación: Las variedades afines de W en V constituyen una partición de V en conjuntos mutuamente disjuntos. Esto es:

i) Dos variedades afines cualesquiera $u + W$ y $v + W$ son idénticas o disjuntas.

ii) Cada $v \in V$ pertenece a una variedad afín; de hecho, $v \in v + W$.

Además, $u + W = v + W$ si y sólo si $u - v \in W$, y así $(v + w) + W = v + W$ para todo $w \in W$.

Sea $v \in V$. Dado que $0 \in W$, tenemos $v = v + 0 \in v + W$, lo que demuestra ii).

Supongamos ahora que las variedades afines $u + W$ y $v + W$ no son disjuntas; digamos que el vector x pertenece simultáneamente a $u + W$ y $v + W$. En tal caso, $u - x \in W$ y $x - v \in W$. La demostración de i) quedará completa si probamos que $u + W = v + W$. Sea $u + w_0$ un elemento arbitrario en la variedad afín $u + W$. Como $u - x$, $x - v$ y w_0 pertenecen a W,

$$(u + w_0) - v = (u - x) + (x - v) + w_0 \in W$$

De este modo, $u + w_0 \in v + W$ y por consiguiente la variedad afín $u + W$ está contenida en la variedad afín $v + W$. Similarmente, $v + W$ está contenida en $u + W$ y así $u + W = v + W$.

La última afirmación se sigue del hecho de que $u + W = v + W$ si y sólo si $u \in v + W$, y, de acuerdo con el Problema 11.21, esto es equivalente a $u - v \in W$.

11.23. Sea W el espacio solución de la ecuación homogénea $2x + 3y + 4z = 0$. Describir las variedades afines de W en $\mathbf{R}^3$.

W es un plano por el origen $O = (0, 0, 0)$, y sus variedades afines son los planos paralelos como se muestra en la Figura 11-3. Equivalentemente, las variedades afines de W son los conjuntos solución de la familia de ecuaciones

$$2x + 3y + 4z = k \qquad k \in \mathbf{R}$$

En particular, la variedad afín $v + W$, donde $v = (a, b, c)$, es el conjunto solución de la ecuación lineal

$$2x + 3y + 4z = 2a + 3b + 4c \qquad \text{o} \qquad 2(x - a) + 3(y - b) + 4(z - c) = 0$$

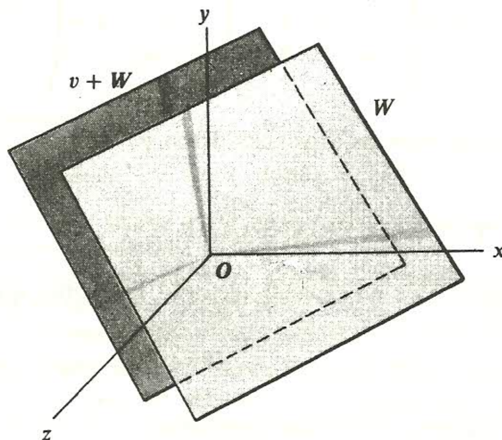


Figura 11-3.

11.24. Supóngase que W es un subespacio de un espacio vectorial V. Mostrar que las operaciones del Teorema 11.15 están bien definidas; a saber, mostrar que si $u + W = u' + W$ y $v + W = v' + W$, se cumple:

a) $(u + v) + W = (u' + v') + W$ y *b*) $ku + W = ku' + W$ para cualquier $k \in K$.

a) Dado que $u + W = u' + W$ y $v + W = v' + W$, tanto $u - u'$ como $v - v'$ pertenecen a W. Pero entonces $(u + v) - (u' + v') = (u - u') + (v - v') \in W$. Por tanto, $(u + v) + W = (u' + v') + W$.

b) Asimismo, como $u - u' \in W$ implica $k(u - u') \in W$, $ku - ku' = k(u - u') \in W$; en consecuencia, $ku + W = ku' + W$.

11.25. Sea V un espacio vectorial y W un subespacio de V. Probar que la aplicación natural $\eta : V \to V/W$, definida por $\eta(v) = v + W$, es lineal.

Para todo par $u, v \in V$ y todo $k \in K$ tenemos

$$\eta(u + v) = u + v + W = u + W + v + W = \eta(u) + \eta(v)$$

y

$$\eta(kv) = kv + W = k(v + W) = k\eta(v)$$

En consecuencia, η es lineal.

11.26. Sea W un subespacio de un espacio vectorial V. Supóngase que $\{w_1, \ldots, w_r\}$ es una base de W y que el conjunto de variedades afines $\{\bar{v}_1, \ldots, \bar{v}_s\}$, donde $\bar{v}_j = v_j + W$, es una base del espacio cociente. Probar que B es una base de V, siendo $B = \{v_1, \ldots, v_s, w_1, \ldots, w_r\}$. De esta manera, $\dim V = \dim W + \dim (V/W)$.

Supongamos $u \in V$. Dado que $\{\bar{v}_j\}$ es una base de V/W,

$$\bar{u} = u + W = a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_s\bar{v}_s$$

Por consiguiente, $u = a_1v_1 + \cdots + a_sv_s + w$, donde $w \in W$. Dado que $\{w_i\}$ es una base de W,

$$u = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_rw_r$$

De acuerdo con ello, B genera V.

Mostremos ahora que B es linealmente independiente. Supongamos

$$c_1v_1 + \cdots + c_sv_s + d_1w_1 + \cdots + d_rw_r = 0 \qquad [1]$$

En ese caso,

$$c_1\bar{v}_1 + \cdots + c_s\bar{v}_s = \bar{0} = W$$

Puesto que $\{\bar{v}_j\}$ es independiente, todas las c son 0. Sustituyendo en [1] encontramos que $d_1w_1 + \cdots + d_rw_r = 0$. Como $\{w_i\}$ es independiente, todas las d son 0. Así B es linealmente independiente y por tanto una base de V.

11.27. Demostrar el Teorema 11.16.

Probemos primero que $\bar{T}$ está bien definido, es decir, si $u + W = v + W$, entonces $\bar{T}(u + W) = \bar{T}(v + W)$. Si $u + W = v + W$, necesariamente $u - v \in W$ y, siendo W T-invariante, $T(u - v) = T(u) - T(v) \in W$. En consecuencia,

$$\bar{T}(u + W) = T(u) + W = T(v) + W = \bar{T}(v + W)$$

como se pedía.

Probemos a continuación que $\bar{T}$ es lineal. Tenemos

$$\begin{aligned}\bar{T}((u + W) + (v + W)) &= \bar{T}(u + v + W) = T(u + v) + W = T(u) + T(v) + W = \\ &= T(u) + W + T(v) + W = \bar{T}(u + W) + \bar{T}(v + W)\end{aligned}$$

y

$$\bar{T}(k(u + W)) = \bar{T}(ku + W) = T(ku) + W = kT(u) + W = k(T(u) + W) = k\bar{T}(u + W)$$

De este modo, $\bar{T}$ es lineal.

Ahora, para toda variedad afín $u + W$ en V/W,

$$\overline{T^2}(u + W) = T^2(u) + W = T(T(u)) + W = \bar{T}(T(u) + W) = \bar{T}(\bar{T}(u + W)) = \bar{T}^2(u + W)$$

De aquí $\overline{T^2} = \bar{T}^2$. Similarmente, $\overline{T^n} = \bar{T}^n$ para todo n. Así para todo polinomio

$$f(t) = a_n t^n + \cdots + a_0 = \sum a_i t^i$$

$$\overline{f(T)}(u + W) = f(T)(u) + W = \sum a_i T^i(u) + W = \sum a_i(T^i(u) + W) =$$
$$= \sum a_i \overline{T^i}(u + W) = \sum a_i \bar{T}^i(u + W) = (\sum a_i \bar{T}^i)(u + W) = f(\bar{T})(u + W)$$

y por tanto $\overline{f(T)} = f(\bar{T})$. En consecuencia, si T es una raíz de $f(t)$, entonces $\overline{f(T)} = \bar{0} = W = f(\bar{T})$, o sea, $\bar{T}$ es también una raíz de $f(t)$. Queda, pues, demostrado el teorema.

11.28. Demostrar el Teorema 11.1.

Se demuestra por inducción en la dimensión de V. Si $\dim V = 1$, toda representación matricial de T es una matriz 1×1, que es triangular.

Supongamos ahora $\dim V = n > 1$ y que el teorema se cumple para polinomios característicos de dimensión menor que n. Dado que el polinomio característico de T se factoriza en polinomios lineales, T tiene al menos un valor propio y por tanto al menos un vector propio no nulo v, digamos $T(v) = a_{11}v$. Sea W el subespacio 1-dimensional generado por v. Hacemos $\bar{V} = V/W$. Entonces (Problema 11.26) $\dim \bar{V} = \dim V - \dim W = n - 1$. Nótese además que W es invariante bajo T. En virtud del Teorema 11.16, T induce un operador lineal $\bar{T}$ en $\bar{V}$ cuyo polinomio mínimo es divisor del de T. Como el polinomio característico de T es un producto de polinomio lineales, también lo es su polinomio mínimo; luego de igual modo lo serán los polinomios característico y mínimo de $\bar{T}$. Así pues, $\bar{V}$ y $\bar{T}$ satisfacen las hipótesis del teorema. Por inducción existe una base $\{\bar{v}_2, \ldots, \bar{v}_n\}$ de $\bar{V}$ tal que

$$\begin{aligned}
\bar{T}(\bar{v}_2) &= a_{22}\bar{v}_2 \\
\bar{T}(\bar{v}_3) &= a_{32}\bar{v}_2 + a_{33}\bar{v}_3 \\
&\cdots\cdots\cdots\cdots \\
\bar{T}(\bar{v}_n) &= a_{n2}\bar{v}_2 + a_{n3}\bar{v}_3 + \cdots + a_{nn}\bar{v}_n
\end{aligned}$$

Sean ahora $v_2, \ldots, v_n$ elementos de V que pertenecen a las variedades afines $\bar{v}_2, \ldots, \bar{v}_n$, respectivamente. En ese caso, $\{v, v_2, \ldots, v_n\}$ es una base de V (Problema 11.26). Puesto que $\bar{T}(\bar{v}_2) = a_{22}\bar{v}_2$, tenemos

$$\bar{T}(\bar{v}_2) - a_{22}\bar{v}_2 = 0 \qquad \text{y por tanto} \qquad T(v_2) - a_{22}v_2 \in W$$

Pero W está generado por v, por lo que $T(v_2) - a_{22}v_2$ es un múltiplo de v, digamos

$$T(v_2) - a_{22}v_2 = a_{21}v \qquad \text{y así} \qquad T(v_2) = a_{21}v + a_{22}v_2$$

De forma similar, para $i = 3, \ldots, n$,

$$T(v_i) - a_{i2}v_2 - a_{i3}v_3 - \cdots - a_{ii}v_i \in W \qquad \text{y así} \qquad T(v_i) = a_{i1}v + a_{i2}v_2 + \cdots + a_{ii}v_i$$

De este modo,

$$\begin{aligned}
T(v) &= a_{11}v \\
T(v_2) &= a_{21}v + a_{22}v_2 \\
&\cdots\cdots\cdots\cdots \\
T(v_n) &= a_{n1}v + a_{n2}v_2 + \cdots + a_{nn}v_n
\end{aligned}$$

y por consiguiente la matriz de T en esta base es triangular.

SUBESPACIOS CICLICOS. FORMA CANONICA RACIONAL

11.29. Demostrar el Teorema 11.12.

i) Por definición de $m_v(t)$, $T^k(v)$ es el primer vector en la sucesión v, $T(v)$, $T^2(v)$, ... que es combinación lineal de los vectores que le preceden; luego el conjunto $B = \{v, T(v), \ldots, T^{k-1}(v)\}$ es linealmente independiente. Ahora sólo tenemos que probar que $Z(v, T) = L(B)$, la envolvente lineal de B. Por lo anterior, $T^k(v) \in L(B)$. Demostramos, por inducción, que $T^n(v) \in L(B)$ para todo n. Supongamos $n > k$ y $T^{n-1}(v) \in L(B)$, esto es, que $T^{n-1}(v)$ es combinación lineal de $v, \ldots, T^{k-1}(v)$. En ese caso, $T^n(v) = T(T^{n-1}(v))$ es combinación lineal de $T(v), \ldots, T^k(v)$. Pero $T^k(v) \in L(B)$, luego $T^n(v) \in L(B)$ para todo n. En consecuencia, $f(T)(v) \in L(B)$ para todo polinomio $f(t)$. Siendo así, $Z(v, T) = L(B)$ y por tanto B es una base como se pretendía.

ii) Supongamos que $m(t) = t^s + b_{s-1}t^{s-1} + \cdots + b_0$ es el polinomio mínimo de T_v. Dado que $v \in Z(v, T)$,

$$0 = m(T_v)(v) = m(T)(v) = T^s(v) + b_{s-1}T^{s-1}(v) + \cdots + b_0 v$$

De este modo, $T^s(v)$ es combinación lineal de v, $T(v)$, ..., $T^{s-1}(v)$ y por consiguiente $k \leq s$. Sin embargo, $m_v(T) = \mathbf{0}$ y así $m_v(T_v) = \mathbf{0}$. Entonces $m(t)$ es divisor de $m_v(t)$ y por esta razón $s \leq k$. De acuerdo con ello, $k = s$ y por tanto $m_v(t) = m(t)$.

iii)

$$\begin{array}{lll} T_v(v) & = & T(v) \\ T_v(T(v)) & = & T^2(v) \\ \multicolumn{3}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ T_v(T^{k-2}(v)) & = & T^{k-1}(v) \\ T_v(T^{k-1}(v)) & = T^k(v) = & -a_0 v - a_1 T(v) - a_2 T^2(v) - \cdots - a_{k-1}T^{k-1}(v) \end{array}$$

Por definición, la matriz de T_v en esta base es la traspuesta de la matriz de coeficientes del sistema de ecuaciones precedente, luego es C, como se pedía.

11.30. Sea $T: V \to V$ lineal. Sea W un subespacio T-invariante de V y $\bar{T}$, el operador inducido en V/W. Demostrar: *a*) El T-aniquilador de $v \in V$ es divisor del polinomio mínimo de T. *b*) El $\bar{T}$-aniquilador de $\bar{v} \in V/W$ es divisor del polinomio mínimo de T.

a) El T-aniquilador de $v \in V$ es el polinomio mínimo de la restricción de T a $Z(v, T)$, y, por tanto, según el Problema 11.6 es divisor del polinomio mínimo de T.

b) El $\bar{T}$-aniquilador de $\bar{v} \in V/W$ es divisor del polinomio mínimo de $\bar{T}$ que a su vez es divisor del polinomio mínimo de T de acuerdo con el Teorema 11.16.

Nota: En caso de que el polinomio mínimo de T sea $f(t)^n$, donde $f(t)$ es un polinomio mínimo irreducible normalizado, el T-aniquilador de $v \in V$ y el $\bar{T}$-aniquilador de $\bar{v} \in V/W$ son de la forma $f(t)^m$ con $m \leq n$.

11.31. Demostrar el Lema 11.13.

Se demuestra por inducción en la dimensión de V. Si $\dim V = 1$, el propio V es T-cíclico y el lema es válido. Supongamos ahora que $\dim V > 1$ y que el lema es válido para aquellos espacios vectoriales de dimensión menor que la de V.

Dado que el polinomio mínimo de T es $f(t)^n$, existe $v_1 \in V$ tal que $f(T)^{n-1}(v_1) \neq 0$; por este motivo, el T-aniquilador de v_1 es $f(t)^n$. Sea $Z_1 = Z(v_1, T)$ y recordemos que Z_1 es T-invariante. Sean $\bar{V} = V/Z_1$ y $\bar{T}$ el operador lineal en $\bar{V}$ inducido por T. En virtud del Teorema 11.6, el

polinomio mínimo de $\bar{T}$ es divisor de $f(t)^n$, luego la hipótesis se cumple para $\bar{V}$ y $\bar{T}$. Consecuentemente, por inducción, $\bar{V}$ es la suma directa de subespacios $\bar{T}$-cíclicos; por ejemplo,

$$\bar{V} = Z(\bar{v}_2, \bar{T}) \oplus \cdots \oplus Z(\bar{v}_r, \bar{T})$$

donde los correspondientes $\bar{T}$-aniquiladores son $f(t)^{n_2}, \ldots, f(t)^{n_r}$, $n \geq n_2 \geq \cdots \geq n_r$.

Afirmamos que existe un vector v_2 en la variedad afín $\bar{v}_2$ cuyo T-aniquilador es $f(t)^{n_2}$, el $\bar{T}$-aniquilador de $\bar{v}_2$. Sea w cualquier vector en $\bar{v}_2$. Entonces $f(T)^{n_2}(w) \in Z_1$. Por consiguiente, existe un polinomio $g(t)$ para el que

$$f(T)^{n_2}(w) = g(T)(v_1) \qquad [1]$$

Siendo $f(t)^n$ el polinomio mínimo de T tenemos por [1],

$$0 = f(T)^n(w) = f(T)^{n-n_2} g(T)(v_1)$$

Pero $f(t)^n$ es el T-aniquilador de v_1, luego $f(t)^n$ es divisor de $f(t)^{n-n_2}g(t)$ y así $g(t) = f(t)^{n_2}h(t)$ para algún polinomio $h(t)$. Hacemos

$$v_2 = w - h(T)(v_1)$$

Como $w - v_2 = h(T)(v_1) \in Z_1$, v_2 también pertenece a la variedad afín $\bar{v}_2$. Así el T-aniquilador de v_2 es múltiplo del $\bar{T}$-aniquilador de $\bar{v}_2$. Por otra parte, por [1],

$$f(T)^{n_2}(v_2) = f(T)^{n_2}(w - h(T)(v_1)) = f(T)^{n_2}(w) - g(T)(v_1) = 0$$

En consecuencia, el T-aniquilador de v_2 es $f(t)^{n_2}$, como se pretendía.

De forma similar, existen vectores $v_3, \ldots, v_r \in V$ tales que $v_i \in \bar{v}_i$ y el T-aniquilador de v_i es $f(t)^{n_i}$, el $\bar{T}$-aniquilador de $\bar{v}_i$. Hacemos

$$Z_2 = Z(v_2, T), \ldots, Z_r = Z(v_r, T)$$

Denotemos por d el grado de $f(t)$, de modo que $f(t)^{n_i}$ tiene grado dn_i. Al ser $f(t)^{n_i}$ el T-aniquilador de v_i y el $\bar{T}$-aniquilador de $\bar{v}_i$ sabemos que

$$\{v_i, T(v_i), \ldots, T^{dn_i - 1}(v_i)\} \qquad \text{y} \qquad \{\bar{v}_i, \bar{T}(\overline{v_i}), \ldots, \bar{T}^{dn_i - 1}(\overline{v_i})\}$$

son bases de $Z(v_i, T)$ y $Z(\bar{v}, \bar{T})$, respectivamente, para $i = 2, \ldots, r$. Pero $\bar{V} = Z(\overline{v_2}, \bar{T}) \oplus \cdots \oplus Z(\overline{v_r}, \bar{T})$; por consiguiente,

$$\{\bar{v}_2, \ldots, \bar{T}^{dn_2 - 1}(\bar{v}_2), \ldots, \bar{v}_r, \ldots, \bar{T}^{dn_r - 1}(\bar{v}_r)\}$$

es una base de $\bar{V}$. Por tanto, según el Problema 11.26 y la relación $\bar{T}^i(\bar{v}) = \overline{T^i(v)}$ (véase el Problema 11.27),

$$\{v_1, \ldots, T^{dn_1 - 1}(v_1), v_2, \ldots, T^{dn_2 - 1}(v_2), \ldots, v_r, \ldots, T^{dn_r - 1}(v_r)\}$$

es una base de V. Por el Teorema 11.4, $V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$, como se pedía.

Queda por probar que los exponentes $n_1, \ldots, n_r$ están unívocamente determinados por T. Puesto que d denota el grado de $f(t)$,

$$\dim V = d(n_1 + \cdots + n_r) \qquad y \qquad \dim Z_i = dn_i \qquad i = 1, \ldots, r$$

Además, si s es cualquier entero positivo, tendremos que (Problema 11.59) $f(T)^s(Z_i)$ es un subespacio cíclico generado por $f(T)^s(v_i)$ y tiene dimensión $d(n_i - s)$ si $n_i > s$ y dimensión 0 si $n_i \leq s$.

Ahora todo vector $v \in V$ puede escribirse de forma única como $v = w_1 + \cdots + w_r$, donde $w_i \in Z_i$. De aquí que cualquier vector en $f(T)^s(V)$ pueda expresarse de forma única como

$$f(T)^s(v) = f(T)^s(w_1) + \cdots + f(T)^s(w_r)$$

donde $f(T)^s(w_i) \in f(T)^s(Z_i)$. Sea t el entero, dependiente de s, para el cual

$$n_1 > s, \ldots, n_t > s, n_{t+1} \geq s$$

En tal caso,
$$f(T)^s(V) = f(T)^s(Z_1) \oplus \cdots \oplus f(T)^s(Z_t)$$

y así
$$\dim (f(T)^s(V)) = d[(n_1 - s) + \cdots + (n_t - s)] \qquad (*)$$

Los números a la izquierda de $(*)$ están unívocamente determinados por T. Tomamos $s = n - 1$ y $(*)$ determina el número de los n_i iguales a n. A continuación tomamos $s = n - 2$ y $(*)$ determina el número de los n_i (si hay alguno) iguales a $n - 1$. Repetimos el proceso hasta que tomemos $s = 0$ y determinemos el número de los n_i iguales a 1. Los n_i están, pues, unívocamente determinados por T y V y queda demostrado el lema.

11.32. Sean V un espacio vectorial de dimensión 7 sobre **R** y $T: V \to V$ un operador lineal con polinomio mínimo $m(t) = (t^2 + 2)(t + 3)^3$. Encontrar todas las formas canónicas racionales posibles para T.

La suma de los grados de las matrices acompañantes debe dar 7. Además, una matriz acompañante debe salir de $t^2 + 2$ y otra de $(t + 3)^3$. De este modo, la forma canónica racional de T es exactamente una de las siguientes sumas directas de matrices acompañantes:

i) $C(t^2 + 2) \oplus C(t^2 + 2) \oplus C((t + 3)^3)$.

ii) $C(t^2 + 2) \oplus C((t + 3)^3) \oplus C((t + 3)^2)$.

iii) $C(t^2 + 2) \oplus C((t + 3)^3) \oplus C(t + 3) \oplus C(t + 3)$.

Esto es,

$$\begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & -2 & & & \\ & & 1 & 0 & & & \\ & & & & 0 & 0 & -27 \\ & & & & 1 & 0 & -27 \\ & & & & 0 & 1 & -9 \end{pmatrix} \qquad \begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & 0 & -9 \\ & & & & & 1 & -6 \end{pmatrix}$$

i) ii)

$$\begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & -3 & \\ & & & & & & -3 \end{pmatrix}$$

iii)

PROYECCIONES

11.33. Supóngase $V = W_1 \oplus \cdots \oplus W_r$. La *proyección* de V en su subespacio W_k es la aplicación $E: V \to V$ definida por $E(v) = w_k$, donde $v = w_1 + \cdots + w_r$, $w_i \in W_i$. Mostrar que: *a*) E es lineal, *b*) $E^2 = E$.

a) Puesto que la suma $v = w_1 + \cdots + w_r$, $w_i \in W$ está unívocamente determinada por v, la aplicación E está bien definida. Supongamos, para $u \in V$, $u = w'_1 + \cdots + w'_r$, $w'_i \in W_i$. En tal caso,

$$v + u = (w_1 + w'_1) + \cdots + (w_r + w'_r) \qquad \text{y} \qquad kv = kw_1 + \cdots + kw_r \qquad kw_i,\ w_i + w'_i \in W_i$$

son las únicas sumas correspondientes a $v + u$ y a kv. De aquí

$$E(v + u) = w_k + w'_k = E(v) + E(u) \qquad \text{y} \qquad E(kv) = kw_k + kE(v)$$

y por consiguiente E es lineal.

b) Tenemos que
$$w_k = 0 + \cdots + 0 + w_k + 0 + \cdots + 0$$
es la única suma correspondiente a $w_k \in W_k$, luego $E(w_k) = w_k$. Entonces, para todo $v \in V$,

$$E^2(v) = E(E(v)) = E(w_k) = w_k = E(v)$$

Siendo así, $E^2 = E$, como se pedía.

11.34. Supóngase que $E: V \to V$ es lineal y que $E^2 = E$. Probar que: *a*) $E(u) = u$ para todo $u \in \text{Im } E$, es decir, la restricción de E a su imagen es la aplicación identidad; *b*) V es la suma directa de la imagen y el núcleo de E: esto es, $V = \text{Im } E \oplus \text{Ker } E$; *c*) E es la proyección de V en Im E, su imagen. Así, por el Problema 11.33, una aplicación lineal $T: V \to V$ es una proyección si y sólo si $T^2 = T$; compárese con el Problema 9.45.

a) Si $u \in \text{Im } E$, existe $v \in V$ para el que $E(v) = u$, luego

$$E(u) = E(E(v)) = E^2(v) = E(v) = u$$

como se pedía.

b) Sea $v \in V$. Podemos escribir v de la forma $v = E(v) + v - E(v)$. Ahora $E(v) \in \text{Im } E$ y, como

$$E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

$v - E(v) \in \text{Ker } E$. En consecuencia, $V = \text{Im } E + \text{Ker } E$.

Supongamos ahora $w \in \text{Im } E \cap \text{Ker } E$. Por i), $E(w) = w$ debido a que $w \in \text{Im } E$. Por otra parte, $E(w) = 0$ porque $w \in \text{Ker } E$. De este modo, $w = 0$ y por tanto $\text{Im } E \cap \text{Ker } E = \{0\}$. Estas dos condiciones implican que V sea la suma directa de la imagen y el núcleo de E.

c) Sea $v \in V$ y supongamos $v = u + w$ con $u \in \text{Im } E$ y $w \in \text{Ker } E$. Nótese que $E(u) = u$, por i), y que $E(w) = 0$ porque $w \in \text{Ker } E$. Por tanto,

$$E(v) = E(u + w) = E(u) + E(w) = u + 0 = u$$

O sea, E es la proyección de V en su imagen.

11.35. Supóngase $V = U \oplus W$ y que $T: V \to V$ es lineal. Probar que U y W son ambos T-invariantes si y sólo si $TE = ET$, donde E es la proyección de V en U.

Observemos que $E(v) \in U$ para todo $v \in V$ y que: i) $E(v) = v$ si y sólo si $v \in U$, ii) $E(v) = 0$ si y sólo si $v \in W$.

Supongamos $ET = TE$. Sea $u \in U$. Puesto que $E(u) = u$,

$$T(u) = T(E(u)) = (TE)(u) = (ET)(u) = E(T(u)) \in U$$

De aquí U es T-invariante. Sea ahora $w \in W$. Dado que $E(w) = 0$,

$$E(T(w)) = (ET)(w) = (TE)(w) = T(E(w)) = T(0) = 0 \quad \text{de modo que} \quad T(w) \in W$$

Por consiguiente, W es también T-invariante.

Recíprocamente, supongamos que U y W son ambos T-invariantes. Sea $v \in V$ y supongamos $v = u + w$ con $u \in U$ y $w \in W$. En ese caso, $T(u) \in U$ y $T(w) \in W$, luego $E(T(u)) = T(u)$ y $E(T(w)) = 0$. Así

$$(ET)(v) = (ET)(u + w) = (ET)(u) + (ET)(w) = E(T(u)) + E(T(w)) = T(u)$$

y

$$(TE)(v) = (TE)(u + w) = T(E(u + w)) = T(u)$$

Esto es, $(ET)(v) = (TE)(v)$ para todo $v \in V$; por esta razón, $ET = TE$, como se pedía.

PROBLEMAS SUPLEMENTARIOS

SUBESPACIOS INVARIANTES

11.36. Supóngase que W es invariante bajo $T: V \to V$. Mostrar que W es invariante bajo $f(T)$ para todo polinomio $f(t)$.

11.37. Probar que todo subespacio es invariante bajo I y bajo $\mathbf{0}$, los operadores identidad y cero.

11.38. Supóngase que W es invariante bajo $T_1: V \to V$ y $T_2: V \to V$. Mostrar que W es también invariante bajo $T_1 + T_2$ y $T_1 T_2$.

11.39. Sea $T: V \to V$ lineal y sea W el espacio propio perteneciente a un valor propio λ de T. Probar que W es T-invariante.

11.40. Sea V un espacio vectorial de dimensión impar (mayor que 1) sobre el cuerpo real $\mathbf{R}$. Probar que todo operador lineal en V tiene un subespacio invariante distinto de V o $\{0\}$.

11.41. Determinar los subespacios invariantes de $A = \begin{pmatrix} 2 & -4 \\ 5 & -2 \end{pmatrix}$ vista como operador lineal en i) $\mathbf{R}^2$, ii) $\mathbf{C}^2$.

11.42. Supóngase dim $V = n$. Mostrar que $T: V \to V$ tiene una representación matricial triangular si y sólo si existen subespacios T-invariantes $W_1 \subset W_2 \subset \cdots \subset W_n = V$ para los que dim $W_k = k$, $k = 1, \ldots, n$.

SUMAS DIRECTAS INVARIANTES

11.43. Se dice que los subespacios $W_1, \ldots, W_r$ son *independientes* si $w_1 + \cdots + w_r = 0$, $w_i \in W_i$, implica que cada $w_i = 0$. Probar que lin $(W_i) = W_1 \oplus \cdots \oplus W_r$ si y sólo si los W_i son independientes. [Aquí lin (W_i) denota la envolvente linea de W_i.]

11.44. Mostrar que $V = W_1 \oplus \cdots \oplus W_r$ si y sólo si: i) $V = \text{lin}(W_i)$, ii) $W_k \cap \text{lin}(W_1, \ldots, W_{k-1}, W_{k+1}, \ldots, W_r) = \{0\}$, $k = 1, \ldots, r$.

11.45. Mostrar que lin $(W_i) = W_1 \oplus \cdots \oplus W_i$ si y sólo si dim lín $(W_i) = \dim W_1 + \cdots + \dim W_r$.

11.46. Supóngase que el polinomio característico de $T: V \to V$ es $\Delta(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_r(t)^{n_r}$, donde los $f_i(t)$ son polinomios irreducibles normalizados distintos. Sea $V = W_1 \oplus \cdots \oplus W_r$ la descomposición primaria de V en subespacios T-invariantes. Probar que $f_i(t)^{n_i}$ es el polinomio característico de la restricción de T a W_i.

OPERADORES NILPOTENTES

11.47. Supóngase que T_1 y T_2 son operadores nilpotentes que conmutan, es decir, $T_1 T_2 = T_2 T_1$. Probar que $T_1 + T_2$ y $T_1 T_2$ son también nilpotentes.

11.48. Supóngase que A es una matriz supertriangular, o sea, que todas sus entradas en y bajo la diagonal principal son 0. Mostrar que A es nilpotente.

11.49. Sea V el espacio vectorial de los polinomios de grado $\leq n$. Probar que el operador de derivación en V es nilpotente de índice $n + 1$.

11.50. Probar que las siguientes matrices nilpotentes de orden n son similares:

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

11.51. Probar que dos matrices nilpotentes de orden 3 son similares si y sólo si tienen el mismo índice de nilpotencia. Mostrar con un ejemplo que la afirmación no es cierta para matrices nilpotentes de orden 4.

FORMA CANONICA DE JORDAN

11.52. Hallar todas las formas canónicas de Jordan posibles para aquellas matrices cuyos polinomios característico $\Delta(t)$ y mínimo $m(t)$ son:

a) $\Delta(t) = (t-2)^4(t-3)^2$, $m(t) = (t-2)^2(t-3)^2$.

b) $\Delta(t) = (t-7)^5$, $m(t) = (t-7)^2$.

c) $\Delta(t) = (t-2)^7$, $m(t) = (t-2)^3$.

d) $\Delta(t) = (t-3)^4(t-5)^4$, $m(t) = (t-3)^2(t-5)^2$.

11.53. Probar que toda matriz compleja es similar a su traspuesta. (*Indicación:* Utilícense su forma canónica de Jordan y el Problema 11.50.)

11.54. Mostrar que todas las matrices complejas $n \times n$ A para las que $A^n = I$ pero $A^k \neq I$ para $k < n$ son similares.

11.55. Supóngase que A es una matriz compleja que tiene sólo valores propios reales. Mostrar que A es similar a una matriz que tiene únicamente entradas reales.

SUBESPACIOS CICLICOS

11.56. Supóngase que $T: V \to V$ es lineal. Demostrar que $Z(v, T)$ es la intersección de todos los subespacios T-invariantes que contienen v.

11.57. Sean $f(t)$ y $g(t)$ los T-aniquiladores de u y v, respectivamente. Mostrar que si $f(t)$ y $g(t)$ son primos entre sí, $f(t)g(t)$ es el T-aniquilador de $u + v$

11.58. Demostrar que $Z(u, T) = Z(v, T)$ si y sólo si $g(T)(u) = v$, donde $g(t)$ y el T-aniquilador de u son primos entre sí.

11.59. Sea $W = Z(v, T)$ y supóngase que el T-aniquilador de v es $f(t)^n$, donde $f(t)$ es un polinomio irreducible normalizado de grado d. Probar que $f(T)^s(W)$ es un subespacio cíclico generado por $f(T)^s(v)$ y tiene dimensión $d(n - s)$ si $n > s$ y 0 si $n \leq s$.

FORMA CANONICA RACIONAL

11.60. Encontrar todas las formas canónicas racionales posibles para:

a) matrices 6×6 con polinomio mínimo $m(t) = (t^2 + 3)(t + 1)^2$.

b) matrices 6×6 con polinomio mínimo $m(t) = (t + 1)^3$.

c) matrices 8×8 con polinomio mínimo $m(t) = (t^2 + 2)^2(t + 3)^2$.

11.61. Sea A una matriz 4×4 con polinomio mínimo $m(t) = (t^2 + 1)(t^2 - 3)$. Hallar la forma canónica racional de A si A es una matriz sobre *a*) el cuerpo racional **Q**, *b*) el cuerpo real **R**, *c*) el cuerpo complejo **C**.

11.62. Hallar la forma canónica racional para el bloque de Jordan $\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.

11.63. Demostrar que el polinomio característico de un operador $T: V \to V$ es un producto de sus divisores elementales.

11.64. Demostrar que dos matrices 3×3 con los mismos polinomios mínimo y característico son similares.

11.65. Denotemos por $C(f(t))$ la matriz acompañante de un polinomio arbitrario $f(t)$. Mostrar que $f(t)$ es el polinomio característico de $C(f(t))$.

PROYECCIONES

11.66. Supóngase $V = W_1 \oplus \cdots \oplus W_r$. Denotemos por E_i la proyección de V en W_i. Demostrar: i) $E_iE_j = 0$, $i \neq j$; ii) $I = E_1 + \cdots + E_r$.

11.67. Sean $E_1, \ldots, E_r$ operadores lineales en V tales que: i) $E_i^2 = E_i$, o sea, los E_i son proyecciones; ii) $E_iE_j = 0$, $i \neq j$; iii) $I = E_1 + \cdots + E_r$. Probar que $V = \text{Im } E_1 \oplus \cdots \oplus \text{Im } E_r$.

11.68. Supóngase que $E: V \to V$ es una proyección, esto es, $E^2 = E$. Demostrar que E tiene una representación matricial de la forma $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, donde r es el rango de E e I_r la matriz identidad r-cuadrada.

11.69. Demostrar que dos proyecciones cualesquiera del mismo rango son similares. (*Indicación:* Hágase uso del Problema 11.68.)

11.70. Supóngase que $E: V \to V$ es una proyección. Demostrar: i) $I - E$ es una proyección y $V = \text{Im } E \oplus \text{Im } (I - E)$, ii) $I + E$ es invertible (si $1 + 1 \neq 0$).

ESPACIOS COCIENTE

11.71. Sea W un subespacio de V. Supóngase que el conjunto de variedades afines $\{v_1 + W, v_2 + W, \ldots, v_n + W\}$ en V/W es linealmente independiente. Mostrar que el conjunto de vectores $\{v_1, v_2, \ldots, v_n\}$ en V es también linealmente independiente.

11.72. Sea W un subespacio de V. Supóngase que el conjunto de vectores $\{u_1, u_2, \ldots, u_n\}$ en V es linealmente independiente y que lin $(u_i) \cap W = \{0\}$. Probar que el conjunto de variedades afines $\{u_1 + W, \ldots, u_n + W\}$ en V/W es también linealmente independiente.

11.73. Supóngase $V = U \oplus W$ y que $\{u_1, \ldots, u_n\}$ es una base de U. Mostrar que $\{u_1 + W, \ldots, u_n + W\}$ es una base del espacio cociente V/W. (Obsérvese que no se ha impuesto condición alguna a las dimensiones de V y W.)

11.74. Sea W el espacio solución de la ecuación lineal

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \qquad a_i \in K$$

y sea $v = (b_1, b_2, \ldots, b_n) \in K^n$. Demostrar que la variedad afín $v + W$ de W en K^n es el conjunto solución de la ecuación lineal

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad \text{donde} \qquad b = a_1b_1 + \cdots + a_nb_n$$

11.75. Sean V el espacio vectorial de los polinomios sobre $\mathbf{R}$ y W el subespacio de los polinomios divisibles por t^4, esto es, de la forma $a_0t^4 + a_1t^5 + \cdots + a_{n-4}t^n$. Probar que el espacio cociente V/W es de dimensión 4.

11.76. Sean U y W subespacios de V tales que $W \subseteq U \subseteq V$. Nótese que cualquier variedad afín $u + W$ de W en U puede verse también como variedad afín de W en V, puesto que $u \in U$ implica $u \in V$. Demostrar que: i) U/W es un subespacio de V/W, ii) $\dim (V/W) - \dim (U/W) = \dim V/U$.

11.77. Sean U y W subespacios de V. Probar que las variedades afines de $U \cap W$ en V pueden obtenerse intersectando cada una de las variedades afines de U en V con cada una de las de W en V:

$$V/(U \cap W) = \{(v + U) \cap (v' + W): v, v' \in V\}$$

11.78. Sea $T: V \to V'$ lineal con núcleo W e imagen U. Mostrar que el espacio cociente V/W es isomorfo a U bajo la aplicación $\theta: V/W \to U$ definida por $\theta(v + W) = T(v)$. Probar además que $T = i \circ \theta \circ \eta$,

donde $\eta: V \to V/W$ es la aplicación natural de V en V/W, es decir, $\eta(v) = v + W$, e $i: U \to V'$ es la aplicación de inclusión, es decir, $i(u) = u$. (Véase la Figura 11-4.)

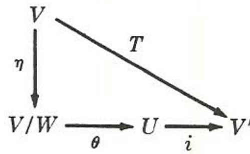


Figura 11-4.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

11.41. *a*) $\mathbf{R}^2$ y $\{0\}$, *b*) $\mathbf{C}^2$, $\{0\}$, $W_1 = L((2, 1 - 2i))$, $W_2 = L((2, 1 + 2i))$.

11.52. *a*)

$$\begin{pmatrix} 2 & 1 & & & & \\ & 2 & & & & \\ & & 2 & 1 & & \\ & & & 2 & & \\ & & & & 3 & 1 \\ & & & & & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & & & & \\ & 2 & & & & \\ & & 2 & & & \\ & & & 2 & & \\ & & & & 3 & 1 \\ & & & & & 3 \end{pmatrix}$$

b)

$$\begin{pmatrix} 7 & 1 & & & \\ & 7 & & & \\ & & 7 & 1 & \\ & & & 7 & \\ & & & & 7 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & & & \\ & 7 & & & \\ & & 7 & & \\ & & & 7 & \\ & & & & 7 \end{pmatrix}$$

c)

$$\begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & 1 & \\ & & & & & 2 & \\ & & & & & & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & & \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & & \\ & & & & & 2 & \\ & & & & & & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ & & & 2 & & & \\ & & & & 2 & & \\ & & & & & 2 & \\ & & & & & & 2 \end{pmatrix}$$

d)

$$\begin{pmatrix} 3 & 1 & & & & & & \\ & 3 & & & & & & \\ & & 3 & 1 & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & & 5 & & \\ & & & & & & 5 & 1 \\ & & & & & & & 5 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & & & & & & \\ & 3 & & & & & & \\ & & 3 & 1 & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & & 5 & & \\ & & & & & & 5 & \\ & & & & & & & 5 \end{pmatrix},$$

$$\begin{pmatrix} 3 & 1 & & & & & & \\ & 3 & & & & & & \\ & & 3 & & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & & 5 & & \\ & & & & & & 5 & 1 \\ & & & & & & & 5 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & & & & & & \\ & 3 & & & & & & \\ & & 3 & & & & & \\ & & & 3 & & & & \\ & & & & 5 & 1 & & \\ & & & & & 5 & & \\ & & & & & & 5 & \\ & & & & & & & 5 \end{pmatrix}$$

11.60. *a)*

$$\begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & -3 & & \\ & & 1 & 0 & & \\ & & & & 0 & -1 \\ & & & & 1 & -2 \end{pmatrix}, \begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & -1 & & \\ & & 1 & -2 & & \\ & & & & 0 & -1 \\ & & & & 1 & -2 \end{pmatrix}, \begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & 1 & & \\ & & 1 & -2 & & \\ & & & & -1 & \\ & & & & & -1 \end{pmatrix}$$

b)

$$\begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & 0 & 0 & -1 \\ & & & 1 & 0 & -3 \\ & & & 0 & 1 & -3 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & 0 & -1 & \\ & & & 1 & -2 & \\ & & & & & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & -1 & & \\ & & & & -1 & \\ & & & & & -1 \end{pmatrix}$$

c)

$$\begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -2 & & \\ & & & & 1 & 0 & & \\ & & & & & & 0 & -9 \\ & & & & & & 1 & -6 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -9 & & \\ & & & & 1 & -6 & & \\ & & & & & & 0 & -9 \\ & & & & & & 1 & -6 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -9 & & \\ & & & & 1 & -6 & & \\ & & & & & & -3 & \\ & & & & & & & -3 \end{pmatrix}$$

11.61. *a*) $\begin{pmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & 0 & \sqrt{3} \\ & & 1 & 0 \end{pmatrix}$ *b*) $\begin{pmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & \sqrt{3} & \\ & & & -\sqrt{3} \end{pmatrix}$ *c*) $\begin{pmatrix} i & & & \\ & -i & & \\ & & \sqrt{3} & \\ & & & -\sqrt{3} \end{pmatrix}$

11.62. $\begin{pmatrix} 0 & 0 & 0 & -\lambda^4 \\ 1 & 0 & 0 & 4\lambda^3 \\ 0 & 1 & 0 & -6\lambda^2 \\ 0 & 0 & 1 & 4\lambda \end{pmatrix}$

CAPITULO 12

Funcionales lineales y espacio dual

12.1. INTRODUCCION

En este capítulo estudiaremos las aplicaciones lineales de un espacio vectorial V en su cuerpo de escalares K. (A menos que se establezca o sobrentienda lo contrario, veremos K como un espacio vectorial sobre sí mismo.) Naturalmente, todos los teoremas y resultados para aplicaciones lineales arbitrarias en V rigen en este caso especial. No obstante, trataremos estas aplicaciones por separado debido a su importancia fundamental y a que la particular relación entre V y K da lugar a nuevas nociones y resultados, no aplicables en el caso general.

12.2. FUNCIONALES LINEALES Y ESPACIO DUAL

Sea V un espacio vectorial sobre un cuerpo K. Una aplicación $\phi: V \to K$ se denomina *funcional lineal* (o *forma lineal*) si para todos los $u, v \in V$ y $a, b \in K$,

$$\phi(au + bv) = a\phi(u) + b\phi(v)$$

En otras palabras, un funcional lineal en V es una aplicación lineal de V en K.

EJEMPLO 12.1

a) Sea $\pi_i: K^n \to K$ la proyección i-ésima, es decir, $\pi_i(a_1, a_2, \ldots, a_n) = a_i$. Entonces π_i es lineal y por tanto es un funcional lineal en K^n.

b) Sean V el espacio vectorial de los polinomios en t sobre $\mathbf{R}$ y $\mathbf{J}: V \to \mathbf{R}$ el operador de integración definido por $\mathbf{J}(p(t)) = \int_0^1 p(t)\,dt$. Recordemos que $\mathbf{J}$ es lineal; por consiguiente, es un funcional lineal en V.

c) Sean V el espacio vectorial de las matrices n-cuadradas sobre K y $T: V \to K$ la aplicación traza

$$T(A) = a_{11} + a_{22} + \cdots + a_{nn} \qquad \text{donde } A = (a_{ij})$$

Esto es, T asigna a cada matriz A la suma de sus elementos diagonales. Esta aplicación es lineal (Problema 12.24), de modo que define un funcional lineal en V.

En virtud del Teorema 9.10, el conjunto de los funcionales lineales en un espacio vectorial V sobre un cuerpo K es asimismo un espacio vectorial sobre K con suma y producto por un escalar definidos según

$$(\phi + \sigma)(v) = \phi(v) + \sigma(v) \qquad \text{y} \qquad (k\phi)(v) = k\phi(v)$$

donde ϕ y σ son funcionales lineales en V y $k \in K$. Este espacio se llama el *espacio dual* de V y se denota por V^*.

EJEMPLO 12.2. Sea $V = K^n$ el espacio vectorial de n-plas que escribimos como vectores columna. El espacio dual V^* puede identificarse con el espacio de vectores fila. En concreto, cualquier funcional lineal ϕ de V^* tiene la representación

$$\phi(x_1, \ldots, x_n) = (a_1, a_2, \ldots, a_n)\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

o simplemente

$$\phi(x_1, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

Históricamente la expresión anterior recibía el nombre de *forma lineal.*

12.3. BASE DUAL

Supongamos que V es un espacio vectorial de dimensión n sobre K. De acuerdo con el Teorema 9.11, la dimensión del espacio dual V^* es también n (ya que K es dimensión 1 sobre sí mismo). De hecho, cada base de V determina una de V^* como sigue (véase la demostración en el Problema 12.3):

Teorema 12.1: Supongamos que $\{v_1, \ldots, v_n\}$ es una base de V sobre K. Sean $\phi_1, \ldots, \phi_n \in V^*$ los funcionales lineales definidos mediante

$$\phi_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En tal caso, $\{\phi_1, \ldots, \phi_n\}$ es una base de V^*.

La base $\{\phi_i\}$ precedente se denomina la base *dual* a $\{v_i\}$ o la *base dual*. La fórmula anterior, que utiliza la delta de Kronecker δ_{ij}, es una forma abreviada de escribir

$$\phi_1(v_1) = 1,\ \phi_1(v_2) = 0,\ \phi_1(v_3) = 0, \ldots, \phi_1(v_n) = 0$$
$$\phi_2(v_1) = 0,\ \phi_2(v_2) = 1,\ \phi_2(v_3) = 0, \ldots, \phi_2(v_n) = 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\phi_n(v_1) = 0,\ \phi_n(v_2) = 0, \ldots, \phi_n(v_{n-1}) = 0,\ \phi_n(v_n) = 1$$

Por el Teorema 9.2, estas aplicaciones lineales ϕ_i son únicas y están bien definidas.

EJEMPLO 12.3. Consideremos la base de $\mathbf{R}^2 : \{v_1 = (2, 1), v_2 = (3, 1)\}$. Hallemos la base dual $\{\phi_1, \phi_2\}$.
Buscamos funcionales lineales $\phi_1(x, y) = ax + by$ y $\phi_2(x, y) = cx + dy$ tales que

$$\phi_1(v_1) = 1 \qquad \phi_1(v_2) = 0 \qquad \phi_2(v_1) = 0 \qquad \phi_2(v_2) = 1$$

Así

$$\left.\begin{array}{l}\phi_1(v_1) = \phi_1(2, 1) = 2a + b = 1\\ \phi_1(v_2) = \phi_1(3, 1) = 3a + b = 0\end{array}\right\} \quad \text{o} \quad a = -1, b = 3$$

$$\left.\begin{array}{l}\phi_2(v_1) = \phi_2(2, 1) = 2c + d = 0\\ \phi_2(v_2) = \phi_2(3, 1) = 3c + d = 1\end{array}\right\} \quad \text{o} \quad c = 1, d = -2$$

Por tanto, la base dual es $\{\phi_1(x, y) = -x + 3y, \phi_2(x, y) = x - 2y\}$.

Los próximos teoremas proporcionan relaciones entre las bases y sus duales.

Teorema 12.2: Sean $\{v_1, \ldots, v_n\}$ una base de V y $\{\phi_1, \ldots, \phi_n\}$ la base dual de V^*. Para todo vector

$$u = \phi_1(u)v_1 + \phi_2(u)v_2 + \cdots + \phi_n(u)v_n \qquad [12.1]$$

y para todo funcional lineal $\sigma \in V^*$,

$$\sigma = \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n \qquad [12.2]$$

Teorema 12.3: Sean $\{v_1, \ldots, v_n\}$ y $\{w_1, \ldots, w_n\}$ bases de V y $\{\phi_1, \ldots, \phi_n\}$ y $\{\sigma_1, \ldots, \sigma_n\}$ las bases de V^* duales a $\{v_i\}$ y $\{w_i\}$, respectivamente. Supongamos que P es la matriz de cambio de base de $\{v_i\}$ a $\{w_i\}$. Entonces $(P^{-1})^T$ es la matriz de cambio de base de $\{\phi_i\}$ a $\{\sigma_i\}$.

12.4. ESPACIO SEGUNDO DUAL

Repetimos que todo espacio vectorial V tiene un espacio dual V^* constituido por los funcionales lineales en V. Así pues, V^* tiene a su vez un espacio dual V^{**}, llamado el *segundo dual* de V, que consiste en todos los funcionales lineales en V^*.

Mostramos ahora que cada $v \in V$ determina un elemento específico $\hat{v} \in V^{**}$. En primer lugar, para todo $\phi \in V^*$ definimos

$$\hat{v}(\phi) = \phi(v)$$

Queda probar que la aplicación $\hat{v}: V^* \to K$ es lineal. Para todo par de escalares $a, b \in K$ y todo par de funcionales lineales $\phi, \sigma \in V^*$ tenemos

$$\hat{v}(a\phi + b\sigma) = (a\phi + b\sigma)(v) = a\phi(v) + b\sigma(v) = a\hat{v}(\phi) + b\hat{v}(\sigma)$$

O sea, $\hat{v}$ es lineal y por tanto $\hat{v} \in V^{**}$. Es válido el teorema que ahora enunciamos, demostrado en el Problema 12.7.

Teorema 12.4: Si V tiene dimensión finita, la aplicación $v \mapsto \hat{v}$ es un isomorfismo de V sobre V^{**}.

La aplicación $v \mapsto \hat{v}$ precedente se conoce como la *aplicación natural* de V en V^{**}. Hacemos énfasis en que esta aplicación nunca es suprayectiva si V no es de dimensión finita. No obstante, siempre es lineal y, más aún, siempre es inyectiva.

Supongamos ahora que V tiene dimensión finita. Según el Teorema 12.4, la aplicación natural determina un isomorfismo entre V y V^{**}. A menos que se especifique lo contrario, identificaremos V con V^{**} mediante esta aplicación. En consecuencia, veremos V como el espacio de funcionales lineales en V^* y escribiremos $V^{**} = V$. Hacemos notar que si $\{\phi_i\}$ es la base de V^* dual a la base $\{v_i\}$ de V, $\{v_i\}$ es la base de $V = V^{**}$ que es dual a $\{\phi_i\}$.

12.5. ANIQUILADORES

Sea W un subconjunto (no necesariamente un subespacio) de un espacio vectorial V. Un funcional lineal $\phi \in V^*$ se llama *un aniquilador* de W si $\phi(w) = 0$ para todo $w \in W$, es decir, si $\phi(W) = \{0\}$. Probemos que el conjunto de todas las aplicaciones tales, denotado por W^0 y denominado *el aniquilador* de W, es un subespacio de V^*. Claramente, $0 \in W^0$. Supongamos ahora $\phi, \sigma \in W^0$. En ese caso, para todo par de escalares $a, b \in K$ y todo $w \in W$,

$$(a\phi + b\sigma)(w) = a\phi(w) + b\sigma(w) = a0 + b0 = 0$$

De este modo, $a\phi + b\sigma \in W^0$ y W^0 es un subespacio de V^*.

En caso de ser W un subespacio de V, disponemos de la siguiente relación entre W y su aniquilador W^0. (Véase la demostración en el Problema 12.11.)

Teorema 12.5: Supongamos que V tiene dimensión finita y que W es un subespacio de V. Entonces:

$$\text{i)} \quad \dim W + \dim W^0 = \dim V \qquad \text{y} \qquad \text{ii)} \quad W^{00} = W$$

Aquí $W^{00} = \{v \in V : \phi(v) = 0 \text{ para todo } \phi \in W^0\}$ o, equivalentemente, $W^{00} = (W^0)^0$, donde W^{00} se ve como subespacio de V bajo la identificación de V y V^{**}.

El concepto de aniquilador nos permite dar otra interpretación a un sistema homogéneo de ecuaciones lineales

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array} \qquad (*)$$

Aquí cada fila $(a_{i1}, a_{i2}, \ldots, a_{in})$ de la matriz de coeficientes $A = (a_{ij})$ se ve como un elemento de K^n, y cada vector solución $\phi = (x_1, x_2, \ldots, x_n)$ como un elemento del espacio dual. En este contexto, el espacio solución S de $(*)$ es el aniquilador de las filas de A y por tanto del espacio fila de A. En consecuencia, haciendo uso del Teorema 12.5, obtenemos otra vez el resultado fundamental que enseguida escribimos sobre la dimensión del espacio solución de un sistema homogéneo de ecuaciones lineales:

$$\dim S = \dim K^n - \dim (\text{espacio fila de } A) = n - \text{rango } A$$

12.6. TRASPUESTA DE UNA APLICACION LINEAL

Sea $T: V \to U$ una aplicación lineal arbitraria de un espacio vectorial V en otro U. Para todo funcional lineal $\phi \in U^*$, la composición $\phi \circ T$ es una aplicación lineal de V en K:

$$V \xrightarrow{T} U \xrightarrow{\phi} K \qquad (\phi \circ T : V \to K)$$

Esto es, $\phi \circ T \in V^*$. Siendo así, la correspondencia

$$\phi \mapsto \phi \circ T$$

es una aplicación de U^* en V^*; la denotaremos por T^t y la llamaremos la traspuesta de T. Dicho de otro modo, $T^t: U^* \to V^*$ se define según

$$T^t(\phi) = \phi \circ T$$

Así $(T^t(\phi))(v) = \phi(T(v))$ para todo $v \in V$.

Teorema 12.6: La aplicación traspuesta T^t anteriormente definida es lineal.

Demostración. Para todo par de escalares $a, b \in K$ y todo par de funcionales lineales $\phi, \sigma \in U^*$,

$$T^t(a\phi + b\sigma) = (a\phi + b\sigma) \circ T = a(\phi \circ T) + b(\sigma \circ T) = aT^t(\phi) + bT^t(\sigma)$$

O sea, T^t es lineal como se pretendía.

Subrayamos que si T es una aplicación lineal de V en U, T^t es una aplicación lineal de U^* en V^*:

$$V \xrightarrow{T} U \qquad V^* \xleftarrow{T^t} U^*$$

El nombre «traspuesta» dado a la aplicación T^t deriva sin duda del siguiente teorema, demostrado en el Problema 12.16.

Teorema 12.7: Sea $T: V \to U$ lineal y sea A la representación matricial de T relativa a las bases $\{v_i\}$ de V y $\{u_i\}$ de U. La matriz traspuesta A^T es la representación matricial de $T^t: U^* \to V^*$ respecto a las bases duales a $\{u_i\}$ y $\{v_i\}$.

PROBLEMAS RESUELTOS

ESPACIOS DUALES Y BASES

12.1. Considérese la base de $\mathbf{R}^3$: $\{v_1 = (1, -1, 3), v_2 = (0, 1, -1), v_3 = (0, 3, -2)\}$. Hallar la base dual $\{\phi_1, \phi_2, \phi_3\}$.

Buscamos funcionales lineales

$$\phi_1(x, y, z) = a_1x + a_2y + a_3z \qquad \phi_2(x, y, z) = b_1x + b_2y + b_3z \qquad \phi_3(x, y, z) = c_1x + c_2y + c_3z$$

tales que

$$\begin{array}{lll} \phi_1(v_1) = 1 & \phi_1(v_2) = 0 & \phi_1(v_3) = 0 \\ \phi_2(v_1) = 0 & \phi_2(v_2) = 1 & \phi_2(v_3) = 0 \\ \phi_3(v_1) = 0 & \phi_3(v_2) = 0 & \phi_3(v_3) = 1 \end{array}$$

Hallamos ϕ_1 como sigue:

$$\begin{aligned} \phi_1(v_1) &= \phi_1(1, -1, 3) = a_1 - a_2 + 3a_3 = 1 \\ \phi_1(v_2) &= \phi_1(0, 1, -1) = \quad a_2 - \ a_3 = 0 \\ \phi_1(v_3) &= \phi_1(0, 3, -2) = \quad 3a_2 - 2a_3 = 0 \end{aligned}$$

Resolviendo el sistema de ecuaciones obtenemos $a_1 = 1$, $a_2 = 0$, $a_3 = 0$. De este modo, $\phi_1(x, y, z) = x$.

A continuación hallamos ϕ_2:

$$\begin{aligned} \phi_2(v_1) &= \phi_2(1, -1, 3) = b_1 - b_2 + 3b_3 = 0 \\ \phi_2(v_2) &= \phi_2(0, 1, -1) = \quad b_2 - \ b_3 = 1 \\ \phi_2(v_3) &= \phi_2(0, 3, -2) = \quad 3b_2 - 2b_3 = 0 \end{aligned}$$

Resolviendo el sistema obtenemos $b_1 = 7$, $b_2 = -2$, $b_3 = -3$. Por tanto, $\phi_2(x, y, z) = 7x - 2y - 3z$.

Finalmente hallamos ϕ_3:

$$\begin{aligned} \phi_3(v_1) &= \phi_3(1, -1, 3) = c_1 - c_2 + 3c_3 = 0 \\ \phi_3(v_2) &= \phi_3(0, 1, -1) = \quad c_2 - \ c_3 = 0 \\ \phi_3(v_3) &= \phi_3(0, 3, -2) = \quad 3c_2 - 2c_3 = 1 \end{aligned}$$

Resolviendo el sistema obtenemos $c_1 = -2$, $c_2 = 1$, $c_3 = 1$. Así $\phi_3(x, y, z) = -2x + y + z$.

12.2. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$ de grado $\leqslant 1$, es decir,

$$V = \{a + bt : a, b \in \mathbf{R}\}$$

Defínanse $\phi_1 : V \to \mathbf{R}$ y $\phi_2 : V \to \mathbf{R}$ según

$$\phi_1(f(t)) = \int_0^1 f(t)\, dt \qquad \text{y} \qquad \phi_2(f(t)) = \int_0^2 f(t)\, dt$$

(Hacemos notar que ϕ_1 y ϕ_2 son lineales, por lo que pertenecen al espacio dual V^*.) Encontrar la base $\{v_1, v_2\}$ de V que es dual a $\{\phi_1, \phi_2\}$.

Sean $v_1 = a + bt$ y $v_2 = c + dt$. Por definición de la base dual,

$$\phi_1(v_1) = 1,\ \phi_2(v_1) = 0 \qquad \text{y} \qquad \phi_1(v_2) = 0,\ \phi_2(v_2) = 1$$

Así

$$\left.\begin{aligned} \phi_1(v_1) &= \int_0^1 (a + bt)\,dt = a + \tfrac{1}{2}b = 1 \\ \phi_2(v_1) &= \int_0^2 (a + bt)\,dt = 2a + 2b = 0 \end{aligned}\right\} \quad \text{o} \quad a = 2,\ b = -2$$

$$\left.\begin{aligned} \phi_1(v_2) &= \int_0^1 (c + dt)\,dt = c + \tfrac{1}{2}d = 0 \\ \phi_2(v_2) &= \int_0^2 (c + dt)\,dt = 2c + 2d = 1 \end{aligned}\right\} \quad \text{o} \quad c = -1/2,\ d = 1$$

En otras palabras, $\{2 - 2t, -\frac{1}{2} + t\}$ es la base de V que es dual a $\{\phi_1, \phi_2\}$.

12.3. Demostrar el Teorema 12.1.

Empezamos probando que $\{\phi_1, \ldots, \phi_n\}$ genera V^*. Sea ϕ un elemento arbitrario de V^* y supongamos

$$\phi(v_1) = k_1,\ \phi(v_2) = k_2,\ \ldots,\ \phi(v_n) = k_n$$

Tomemos $\sigma = k_1\phi_1 + \cdots + k_n\phi_n$. Entonces

$$\begin{aligned} \sigma(v_1) &= (k_1\phi_1 + \cdots + k_n\phi_n)(v_1) = k_1\phi_1(v_1) + k_2\phi_2(v_1) + \cdots + k_n\phi_n(v_1) = \\ &= k_1 \cdot 1 + k_2 \cdot 0 + \cdots + k_n \cdot 0 = k_1 \end{aligned}$$

De forma similar, para $i = 2, \ldots, n$,

$$\sigma(v_i) = (k_1\phi_1 + \cdots + k_n\phi_n)(v_i) = k_1\phi_1(v_i) + \cdots + k_i\phi_i(v_i) + \cdots + k_n\phi_n(v_i) = k_i$$

De este modo, $\phi(v_i) = \sigma(v_i)$ para $i = 1, \ldots, n$. Dado que ϕ y σ coinciden sobre los vectores de la base, $\phi = \sigma = k_1\phi_1 + \cdots + k_n\phi_n$. De acuerdo con ello, $\{\phi_1, \ldots, \phi_n\}$ genera V^*.

Falta probar que $\{\phi_1, \ldots, \phi_n\}$ es linealmente independiente. Supongamos

$$a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n = 0$$

Aplicando ambos miembros a v_1,

$$\begin{aligned} 0 = \mathbf{0}(v_1) &= (a_1\phi_1 + \cdots + a_n\phi_n)(v_1) = a_1\phi_1(v_1) + a_2\phi_2(v_1) + \cdots + a_n\phi_n(v_1) = \\ &= a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_n \cdot 0 = a_1 \end{aligned}$$

Similarmente, para $i = 2, \ldots, n$,

$$0 = \mathbf{0}(v_i) = (a_1\phi_1 + \cdots + a_n\phi_n)(v_i) = a_1\phi_1(v_i) + \cdots + a_i\phi_i(v_i) + \cdots + a_n\phi_n(v_i) = a_i$$

Esto es, $a_1 = 0, \ldots, a_n = 0$. Por tanto, $\{\phi_1, \ldots, \phi_n\}$ es linealmente independiente, de modo que es una base de V^*.

12.4. Demostrar el Teorema 12.2.

Supongamos

$$u = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \qquad [1]$$

En tal caso,

$$\phi_1(u) = a_1\phi_1(v_1) + a_2\phi_1(v_2) + \cdots + a_n\phi_1(v_n) = a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_n \cdot 0 = a_1$$

De forma similar, para $i = 2, \ldots, n$,

$$\phi_i(u) = a_1\phi_i(v_1) + \cdots + a_i\phi_i(v_i) + \cdots + a_n\phi_i(v_n) = a_i$$

Esto es, $\phi_1(u) = a_1$, $\phi_2(u) = a_2$, ..., $\phi_n(u) = a_n$. Sustituyendo estos resultados en [1] llegamos a [12.1].

A continuación demostramos [12.2]. Aplicando el funcional lineal σ a ambos miembros de [12.1],

$$\begin{aligned}\sigma(u) &= \phi_1(u)\sigma(v_1) + \phi_2(u)\sigma(v_2) + \cdots + \phi_n(u)\sigma(v_n) = \\ &= \sigma(v_1)\phi_1(u) + \sigma(v_2)\phi_2(u) + \cdots + \sigma(v_n)\phi_n(u) = \\ &= (\sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n)(u)\end{aligned}$$

Puesto que lo anterior se cumple para todo $u \in V$, $\sigma = \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi$, como se pretendía.

12.5. Demostrar el Teorema 12.3.

Supongamos

$$\begin{array}{ll} w_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n & \sigma_1 = b_{11}\phi_1 + b_{12}\phi_2 + \cdots + b_{1n}\phi_n \\ w_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & \sigma_2 = b_{21}\phi_1 + b_{22}\phi_2 + \cdots + b_{2n}\phi_n \\ \cdots\cdots\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ w_n = a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n & \sigma_n = b_{n1}\phi_1 + b_{n2}\phi_2 + \cdots + b_{nn}\phi_n \end{array}$$

donde $P = (a_{ij})$ y $Q = (b_{ij})$. Intentamos demostrar que $Q = (P^{-1})^T$.

Denotemos por R_i la fila i-ésima de Q y por C_j la columna j-ésima de P^T. Entonces

$$R_i = (b_{i1}, b_{i2}, \ldots, b_{in}) \qquad \text{y} \qquad C_j = (a_{j1}, a_{j2}, \ldots, a_{jn})^T$$

Por definición de la base dual,

$$\begin{aligned}\sigma_i(w_j) &= (b_{i1}\phi_1 + b_{i2}\phi_2 + \cdots + b_{in}\phi_n)(a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{jn}v_n) = \\ &= b_{i1}a_{j1} + b_{i2}a_{j2} + \cdots + b_{in}a_{jn} = R_i C_j = \delta_{ij}\end{aligned}$$

donde δ_{ij} es la delta de Kronecker. Así pues,

$$QP^T = \begin{pmatrix} R_1C_1 & R_1C_2 & \ldots & R_nC_n \\ R_2C_1 & R_2C_2 & \ldots & R_2C_n \\ \cdots & \cdots & \cdots & \cdots \\ R_nC_1 & R_nC_2 & \ldots & R_nC_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I$$

y por tanto $Q = (P^T)^{-1} = (P^{-1})^T$, como se pretendía.

12.6. Supóngase que V tiene dimensión finita. Mostrar que si $v \in V$, $v \neq 0$, existe $\phi \in V^*$ tal que $\phi(v) \neq 0$.

Extendemos $\{v\}$ a una base $\{v, v_2, \ldots, v_n\}$ de V. Por el Teorema 9.2 existe una única aplicación lineal $\phi: V \to K$ tal que $\phi(v) = 1$ y $\phi(v_i) = 0$, $i = 2, \ldots, n$. Por consiguiente, ϕ tiene la propiedad deseada.

12.7. Demostrar el Teorema 12.4.

Demostramos primero que la aplicación $v \mapsto \hat{v}$ es lineal, esto es, que para todo par de vectores v, $w \in V$ y todo par de escalares a, $b \in K$, $\widehat{av + bw} = a\hat{v} + b\hat{w}$. Para todo funcional lineal $\phi \in V^*$,

$$\widehat{av + bw}(\phi) = \phi(av + bw) = a\phi(v) + b\phi(w) = a\hat{v}(\phi) + b\hat{w}(\phi) = (a\hat{v} + b\hat{w})(\phi)$$

Como $\widehat{av + bw}(\phi) = (a\hat{v} + b\hat{w})(\phi)$ para todo $\phi \in V^*$, tenemos $\widehat{av + bw} = a\hat{v} + b\hat{w}$. De esta manera, la aplicación $v \mapsto \hat{v}$ es lineal.

Supongamos ahora $v \in V$, $v \neq 0$. Entonces, por el Problema 12.6, existe $\phi \in V^*$ para el cual $\phi(v) \neq 0$. De aquí $\hat{v}(\phi) = \phi(v) \neq 0$, luego $\hat{v} \neq 0$. Como $v \neq 0$ implica $\hat{v} \neq 0$, la aplicación $v \mapsto \hat{v}$ es no singular, luego es un isomorfismo (Teorema 9.9).

Ahora dim V = dim V^* = dim V^{**} porque V tiene dimensión finita. En consecuencia, la aplicación $v \mapsto \hat{v}$ es un isomorfismo de V sobre V^{**}.

ANIQUILADORES

12.8. Probar que si $\phi \in V^*$ aniquila un subconjunto S de V, ϕ necesariamente aniquila la envolvente lineal $L(S)$ de S. Por tanto, $S^0 = (\text{lin}\ (S))^0$.

Supongamos $v \in \text{lin}\ (S)$. En ese caso, existen $w_1, \ldots, w_r \in S$ para los cuales $v = a_1 w_1 + a_2 w_2 + \cdots + a_r w_r$.

$$\phi(v) = a_1\phi(w_1) + a_2\phi(w_2) + \cdots + a_r\phi(w_r) = a_1 0 + a_2 0 + \cdots + a_r 0 = 0$$

Siendo v un elemento arbitrario de lin (S), ϕ aniquila lin (S), como se pretendía demostrar.

12.9. Sea W el subespacio de $\mathbf{R}^4$ generado por $v_1 = (1, 2, -3, 4)$ y $v_2 = (0, 1, 4, -1)$. Encontrar una base del aniquilador de W.

De acuerdo con el Problema 12.8, es suficiente hallar una base del conjunto de funcionales lineales de la forma $\phi(x, y, z, w) = ax + by + cz + dw$ para los que $\phi(v_1) = 0$ y $\phi(v_2) = 0$:

$$\begin{aligned} \phi(1, 2, -3, 4) &= a + 2b - 3c + 4d = 0 \\ \phi(0, 1, 4, -1) &= b + 4c - d = 0 \end{aligned}$$

El sistema de ecuaciones con incógnitas a, b, c, d está en forma escalonada con variables libres c y d.

Tomamos $c = 1$, $d = 0$ para obtener la solución $a = 11$, $b = -4$, $c = 1$, $d = 0$ y por tanto el funcional lineal $\phi_1(x, y, z, w) = 11x - 4y + z$.

Tomamos $c = 0$, $d = -1$ para obtener la solución $a = 6$, $b = -1$, $c = 0$, $d = -1$ y por ende el funcional lineal $\phi_2(x, y, z, w) = 6x - y - w$.

El conjunto de funcionales lineales $\{\phi_1, \phi_2\}$ es una base de W^0, el aniquilador de W.

12.10. Probar que: *a*) para todo subconjunto S de V, $S \subseteq S^{00}$; *b*) si $S_1 \subseteq S_2$, entonces $S_2^0 \subseteq S_1^0$.

a) Sea $v \in S$. Para todo funcional lineal $\phi \in S^0$, $\hat{v}(\phi) = \phi(v) = 0$. De aquí $\hat{v} \in (S^0)^0$. Por tanto, bajo la identificación de V y V^{**}, $v \in S^{00}$. En consecuencia, $S \subseteq S^{00}$.

b) Sea $\phi \in S_2^0$. Entonces $\phi(v) = 0$ para todo $v \in S_2$. Pero $S_1 \subseteq S_2$, luego ϕ aniquila todo elemento de S_1, o sea, $\phi \in S_1$. Por consiguiente, $S_2^0 \subseteq S_1^0$.

12.11. Demostrar el Teorema 12.5.

i) Supongamos dim $V = n$ y dim $W = r \leqslant n$. Queremos probar que dim $W^0 = n - r$. Escogemos una base $\{w_1, \ldots, w_r\}$ de W y la extendemos a la siguiente base de V: $\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$. Consideremos la base dual

$$\{\phi_1, \ldots, \phi_r, \sigma_1, \ldots, \sigma_{n-r}\}$$

Por definición de la base dual, cada una de las σ precedentes aniquila cada w_i, luego $\sigma_1, \ldots, \sigma_{n-r} \in W^0$. Afirmamos que $\{\sigma_j\}$ es una base de W^0. Al ser $\{\sigma_j\}$ parte de una base de V^*, es linealmente independiente

Probamos a continuación que $\{\sigma_j\}$ genera W^0. Sea $\sigma \in W^0$. Según el Teorema 12.2,

$$\begin{aligned}\sigma &= \sigma(w_1)\phi_1 + \cdots + \sigma(w_r)\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} = \\ &= 0\phi_1 + \cdots + 0\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} = \\ &= \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r}\end{aligned}$$

En consecuencia, $\{\sigma_1, \ldots, \sigma_{n-r}\}$ genera W^0 y por tanto es una base de W^0. De acuerdo con ello, como se pedía,

$$\dim W^0 = n - r = \dim V - \dim W.$$

ii) Supongamos dim $V = n$ y dim $W = r$. En tal caso, dim $V^* = n$ y por i) dim $W^0 = n - r$. Así, por i), dim $W^{00} = n - (n - r) = r$; por consiguiente, dim $W =$ dim W^{00}. Según el Problema 12.10, $W \subseteq W^{00}$. En consecuencia, $W = W^{00}$.

12.12. Sean U y W subespacios de V. Demostrar: $(U + W)^0 = U^0 \cap W^0$.

Sea $\phi \in (U + W)^0$. Entonces ϕ aniquila $U + W$, por lo que, en particular, aniquila U y W. Esto es, $\phi \in U^0$ y $\phi \in W^0$, luego $\phi \in U^0 \cap W^0$. Siendo así, $(U + W)^0 \subseteq U^0 \cap W^0$.

Por otra parte, supongamos $\sigma \in U^0 \cap W^0$. En ese caso, σ aniquila U y también W. Si $v \in U + W$, $v = u + w$ con $u \in U$ y $w \in W$. Por tanto, $\sigma(v) = \sigma(u) + \sigma(w) = 0 + 0 = 0$. De este modo, σ aniquila $U + W$, es decir, $\sigma \in (U + W)^0$. En consecuencia, $U^0 + W^0 \subseteq (U + W)^0$.

Ambas relaciones de inclusión nos conducen a la igualdad deseada.

Nota: Obsérvese que no se ha empleado ningún argumento de dimensión en la demostración, luego el resultado es válido para espacios de dimensión finita o infinita.

TRASPUESTA DE UNA APLICACION LINEAL

12.13. Sea ϕ el funcional lineal en $\mathbf{R}^2$ definido por $\phi(x, y) = x - 2y$. Para cada uno de los siguientes operadores lineales T en $\mathbf{R}^2$, hallar $(T^t(\phi))(x, y)$:

a) $T(x, y) = (x, 0)$, *b*) $T(x, y) = (y, x + y)$, *c*) $T(x, y) = (2x - 3y, 5x + 2y)$.

Por definición de la aplicación traspuesta, $T^t(\phi) = \phi \circ T$, esto es, $(T^t(\phi))(v) = \phi(T(v))$ para todo vector v. De donde:

a) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(x, 0) = x$.

b) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(y, x + y) = y - 2(x + y) = -2x - y.$

c) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(2x - 3y, 5x + 2y) = (2x - 3y) - 2(5x + 2y) = -8x - 7y.$

12.14. Sea $T: V \to U$ lineal y sea $T^t: U^* \to V^*$ su traspuesta. Mostrar que el núcleo de T^t es el aniquilador de la imagen de T, o sea, $\text{Ker } T^t = (\text{Im } T)^0$.

Supongamos $\phi \in \text{Ker } T^t$; esto es, $T^t(\phi) = \phi \circ T = \mathbf{0}$. Si $u \in \text{Im } T$, necesariamente $u = T(v)$ para algún $v \in V$, luego

$$\phi(u) = \phi(T(v)) = (\phi \circ T)(v) = \mathbf{0}(v) = 0$$

Tenemos que $\phi(u) = 0$ para todo $u \in \text{Im } T$, luego $\phi \in (\text{Im } T)^0$. Así pues, $\text{Ker } T^t \subseteq (\text{Im } T)^0$.

Por otra parte, supongamos $\sigma \in (\text{Im } T)^0$; es decir, $\sigma(\text{Im } T) = \{0\}$. Entonces, para todo $v \in V$,

$$(T^t(\sigma))(v) = (\sigma \circ T)(v) = \sigma(T(v)) = 0 = \mathbf{0}(v)$$

Tenemos que $(T^t(\sigma))(v) = \mathbf{0}(v)$ para todo $v \in V$, de modo que $T^t(\sigma) = \mathbf{0}$. Por tanto, $\sigma \in \text{Ker } T^t$, de modo que $(\text{Im } T)^0 \subseteq \text{Ker } T^t$.

Ambas relaciones de inclusión nos proporcionan la igualdad requerida.

12.15. Supóngase que V y U tienen dimensión finita y que $T: V \to U$ es lineal. Demostrar: rango T = rango T^t.

Supongamos $\dim V = n$ y $\dim U = m$. Supongamos, asimismo, que rango $T = r$. Según el Teorema 12.5,

$$\dim (\text{Im } T)^0 = \dim U - \dim (\text{Im } T) = m - \text{rango } T = m - r$$

Por el Problema 12.14, $\text{Ker } T^t = (\text{Im } T)^0$. Por consiguiente, nulidad $T^t = m - r$. Se deduce entonces que, como se pretendía,

$$\text{rango } T^t = \dim U^* - \text{nulidad } T^t = m - (m - r) = r = \text{rango } T$$

12.16. Demostrar el Teorema 12.7.

Supongamos

$$\begin{aligned} T(v_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ T(v_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(v_m) &= a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n \end{aligned} \qquad [1]$$

Queremos demostrar que

$$\begin{aligned} T^t(\sigma_1) &= a_{11}\phi_1 + a_{21}\phi_2 + \cdots + a_{m1}\phi_m \\ T^t(\sigma_2) &= a_{12}\phi_1 + a_{22}\phi_2 + \cdots + a_{m2}\phi_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T^t(\sigma_n) &= a_{1n}\phi_1 + a_{2n}\phi_2 + \cdots + a_{mn}\phi_m \end{aligned} \qquad [2]$$

donde $\{\sigma_i\}$ y $\{\phi_j)$ son las bases duales a $\{u_i\}$ y $\{v_j\}$, respectivamente.

Sea $v \in V$ y supongamos $v = k_1v_1 + k_2v_2 + \cdots + k_mv_m$. En tal caso, por [1],

$$\begin{aligned} T(v) &= k_1T(v_1) + k_2T(v_2) + \cdots + k_mT(v_m) = \\ &= k_1(a_{11}u_1 + \cdots + a_{1n}u_n) + k_2(a_{21}u_1 + \cdots + a_{2n}u_n) + \cdots + k_m(a_{m1}u_1 + \cdots + a_{mn}u_n) = \\ &= (k_1a_{11} + k_2a_{21} + \cdots + k_ma_{m1})u_1 + \cdots + (k_1a_{1n} + k_2a_{2n} + \cdots + k_ma_{mn})u_n = \\ &= \sum_{i=1}^{n}(k_1a_{1i} + k_2a_{2i} + \cdots + k_ma_{mi})u_i \end{aligned}$$

Por tanto, para $j = 1, \ldots, n$,

$$(T^t(\sigma_j)(v)) = \sigma_j(T(v)) = \sigma_j\left(\sum_{i=1}^{n}(k_1a_{1i} + k_2a_{2i} + \cdots + k_ma_{mi})u_i\right) =$$
$$= k_1a_{1j} + k_2a_{2j} + \cdots + k_ma_{mj} \qquad [3]$$

Por otra parte, para $j = 1, \ldots, n$,

$$(a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(v) = (a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(k_1v_1 + k_2v_2 + \cdots + k_mv_m) =$$
$$= k_1a_{1j} + k_2a_{2j} + \cdots + k_ma_{mj} \qquad [4]$$

Como $v \in V$ era arbitrario, [3] y [4] implican que

$$T^t(\sigma_j) = a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m \qquad j = 1, \ldots, n$$

que es [2]. Queda, pues, demostrado el teorema.

12.17. Sea A una matriz $m \times n$ sobre un cuerpo K arbitraria. Demostrar que el rango por filas y el rango por columnas de A son iguales.

Sea $T: K^n \to K^m$ la aplicación lineal definida por $T(v) = Av$, donde los elementos de K^n y K^m se escriben como vectores columna. Entonces A es la representación matricial de T relativa a las bases usuales de K^n y K^m y la imagen de T el espacio columna de A. Por consiguiente,

$$\text{rango } T = \text{rango por columnas de } A$$

En virtud del Teorema 12.7, A^T es la representación matricial de T^t respecto a las bases duales, luego

$$\text{rango } T^t = \text{rango por columnas de } A^T = \text{rango por filas de } A$$

Pero por el Problema 12.15, rango T = rango T^t, por lo que el rango por filas y el rango por columnas de A son iguales. (Este resultado se estableció anteriormente como Teorema 5.18 y se demostró de forma directa en el Problema 5.53.)

PROBLEMAS SUPLEMENTARIOS

ESPACIOS DUALES Y BASES DUALES

12.18. Sean $\phi: \mathbf{R}^3 \to \mathbf{R}$ y $\sigma: \mathbf{R}^3 \to \mathbf{R}$ los funcionales lineales definidos por $\phi(x, y, z) = 2x - 3y + z$ y $\sigma(x, y, z) = 4x - 2y + 3z$. Hallar: *a*) $\phi + \sigma$, *b*) 3ϕ, *c*) $2\phi - 5\sigma$.

12.19. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$ de grado $\leqslant 2$. Sean ϕ_1, ϕ_2 y ϕ_3 los funcionales lineales en V definidos por

$$\phi_1(f(t)) = \int_0^1 f(t)\,dt \qquad \phi_2(f(t)) = f'(1) \qquad \phi_3(f(t)) = f(0)$$

Aquí $f(t) = a + bt + ct^2 \in V$ y $f'(t)$ denota la derivada de $f(t)$. Hallar la base $\{f_1(t), f_2(t), f_3(t)\}$ de V que es dual a $\{\phi_1, \phi_2, \phi_3\}$.

12.20. Supóngase que $u, v \in V$ y que $\phi(u) = 0$ implica $\phi(v) = 0$ para todos los $\phi \in V^*$. Mostrar que $v = ku$ para algún escalar k.

12.21. Supóngase que $\phi, \sigma \in V^*$ y que $\phi(v) = 0$ implica $\sigma(v) = 0$ para todos los $v \in V$. Mostrar que $\sigma = k\phi$ para algún escalar k.

12.22. Sea V el espacio vectorial de los polinomios sobre K. Para $a \in K$, defínase $\phi_a : V \to K$ según $\phi_a(f(t)) = f(a)$. Probar que: *a*) ϕ_a es lineal; *b*) si $a \neq b$, necesariamente $\phi_a \neq \phi_b$.

12.23. Sea V el espacio vectorial de los polinomios de grado $\leqslant 2$. Sean $a, b, c \in K$ escalares distintos y ϕ_a, ϕ_b, ϕ_c los funcionales lineales definidos por $\phi_a(f(t)) = f(a)$, $\phi_b(f(t)) = f(b)$, $\phi_c(f(t)) = f(c)$. Probar que $\{\phi_a, \phi_b, \phi_c\}$ es linealmente independiente y encontrar la base $\{f_1(t), f_2(t), f_3(t)\}$ de V que es su dual.

12.24. Sea V el espacio vectorial de las matrices cuadradas de orden n. Sea $T : V \to K$ la aplicación traza: esto es, $T(A) = a_{11} + a_{22} + \cdots + a_{nn}$, donde $A = (a_{ij})$. Probar que T es lineal.

12.25. Sea W un subespacio de V. Para todo funcional lineal ϕ en W, mostrar que existe uno σ en V tal que $\sigma(w) = \phi(w)$ para todo $w \in W$, es decir, ϕ es la restricción de σ a W.

12.26. Sea $\{e_1, \ldots, e_n\}$ la base usual de K^n. Probar que la base dual es $\{\pi_1, \ldots, \pi_n\}$, donde π_i es la proyección i-ésima: esto es, $\pi_i(a_1, \ldots, a_n) = a_i$.

12.27. Sea V un espacio vectorial sobre $\mathbf{R}$. Sean $\phi_1, \phi_2 \in V^*$ y supóngase que $\sigma : V \to R$ definida por $\sigma(v) = \phi_1(v)\phi_2(v)$ también pertenece a V^*. Mostrar que $\phi_1 = \mathbf{0}$ o $\phi_2 = \mathbf{0}$.

ANIQUILADORES

12.28. Sea W el subespacio de $\mathbf{R}^4$ generado por $(1, 2, -3, 4)$, $(1, 3, -2, 6)$ y $(1, 4, -1, 8)$. Hallar una base del aniquilador de W.

12.29. Sea W el subespacio de $\mathbf{R}^3$ generado por $(1, 1, 0)$ y $(0, 1, 1)$. Hallar una base del aniquilador de W.

12.30. Mostrar que, para todo subconjunto S de V, lin $(S) = S^{00}$, siendo lin (S) la envolvente lineal de S.

12.31. Sean U y W subespacios de un espacio vectorial V de dimensión finita. Demostrar:

$$(U \cap W)^0 = U^0 + W^0$$

12.32. Supóngase $V = U \oplus W$. Demostrar que $V^* = U^0 \oplus W^0$.

TRASPUESTA DE UNA APLICACION LINEAL

12.33. Sea ϕ el funcional lineal en $\mathbf{R}^2$ definido por $\phi(x, y) = 3x - 2y$. Para cada una de las siguientes aplicaciones lineales $T : \mathbf{R}^3 \to \mathbf{R}^2$, hallar $(T^t(\phi))(x, y, z)$.

a) $T(x, y, z) = (x + y, y + z)$; *b*) $T(x, y, z) = (x + y + z, 2x - y)$.

12.34. Supóngase que $T_1 : U \to V$ y $T_2 : V \to W$ son lineales. Demostrar que $(T_2 \circ T_1)^t = T_1^t \circ T_2^t$.

12.35. Supóngase que $T : V \to U$ es lineal y que V tiene dimensión finita. Demostrar que Im $T^t = (\text{Ker } T)^0$.

12.36. Supóngase que $T: V \to U$ es lineal y $u \in U$. Demostrar que $u \in \operatorname{Im} T$ o existe $\phi \in V^*$ tal que $T^t(\phi) = 0$ y $\phi(u) = 1$.

12.37. Sea V de dimensión finita. Probar que la aplicación $T \mapsto T^t$ es un isomorfismo de Hom (V, V) sobre Hom (V^*, V^*). (Aquí T es cualquier operador lineal en V.)

PROBLEMAS VARIOS

12.38. Sean V un espacio vectorial sobre $\mathbf{R}$ y $\phi \in V^*$. Se definen:

$$W^+ = \{v \in V: \phi(v) > 0\} \qquad W = \{v \in V: \phi(v) = 0\} \qquad W^- = \{v \in V: \phi(v) < 0\}$$

Demostrar que W^+, W y W^- son convexos (véanse los Problemas 9.90 y 9.91).

12.39. Sea V un espacio vectorial de dimensión finita. Un *hiperplano* H de V puede definirse como el núcleo de un funcional lineal no nulo ϕ en V. [Con esta definición, un hiperplano «pasa por el origen» necesariamente, esto es, contiene el vector cero.] Mostrar que todo subespacio de V es la intersección de un número finito de hiperplanos.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

12.18. *a)* $6x - 5y + 4z$, *b)* $6x - 9y + 3z$, *c)* $-16x + 4y - 13z$.

12.22. *b)* Sea $f(t) = t$. Entonces $\phi_a(f(t)) = a \neq b = \phi_b(f(t))$ y por tanto $\phi_a \neq \phi_b$.

12.23. $\left\{ f_1(t) = \dfrac{t^2 - (b+c)t + bc}{(a-b)(a-c)}, \quad f_2(t) = \dfrac{t^2 - (a+c)t + ac}{(b-a)(b-c)}, \quad f_3(t) = \dfrac{t^2 - (a+b)t + ab}{(c-a)(c-b)} \right\}$

12.28. $\{\phi_1(x, y, z, t) = 5x - y + z,\ \phi_2(x, y, z, t) = 2y - t\}$.

12.29. $\{\phi(x, y, z) = x - y + z\}$.

12.33. *a)* $(T^t(\phi))(x, y, z) = 3x + y - 2z$, *b)* $(T^t(\phi))(x, y, z) = -x + 5y + 3z$.

CAPITULO 13

Formas bilineales, cuadráticas y hermíticas

13.1. INTRODUCCION

En el presente capítulo se generalizan las nociones de aplicaciones lineales y funcionales lineales. Específicamente, introduciremos el concepto de forma bilineal. (En realidad, las aplicaciones multilineales generales ya aparecían en la Sección 7.13.) Estas aplicaciones bilineales dan lugar a su vez a las formas cuadráticas y hermíticas. Aunque las formas cuadráticas surgieron previamente en el contexto de las matrices, este capítulo se trata con independencia de los resultados precedentes. (Así que puede haber cierto solapamiento en la discusión y en algunos ejemplos y problemas.)

13.2. FORMAS BILINEALES

Sea V un espacio vectorial de dimensión finita sobre un cuerpo K. Una *forma bilineal* en V es una aplicación $f: V \times V \to K$ que satisface:

i) $f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$

ii) $f(u, av_1 + bv_2) = af(u, v_1) + bf(u, v_2)$

para todos los $a, b \in K$ y todos los $u_i, v_i \in V$. Expresamos la condición i) diciendo que f es *lineal en la primera variable* y la ii) diciendo que f es *lineal en la segunda variable*.

EJEMPLO 13.1

a) Sean ϕ y σ funcionales lineales arbitrarios en V. Definamos $f: V \times V \to K$ por $f(u, v) = \phi(u)\sigma(v)$. En tal caso, f es bilineal porque ϕ y σ son lineales. (Tal forma bilineal f resulta ser el «producto tensorial» de ϕ y σ y por eso se escribe a veces $f = \phi \otimes \sigma$.)

b) Sea f el producto escalar en $\mathbf{R}^n$; esto es,

$$f(u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

donde $u = (a_i)$ y $v = (b_i)$. En tal caso, f es una forma bilineal en $\mathbf{R}^n$.

c) Sea $A = (a_{ij})$ cualquier matriz $n \times n$ sobre K. La matriz A puede ser identificada con una forma bilineal f en K^n, donde

$$f(X, Y) = X^T A Y = (x_1, x_2, \ldots, x_n) \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix} =$$

$$= \sum_{i,j=1}^{n} a_{ij} x_i y_i = a_{11} x_1 y_1 + a_{12} x_1 y_2 + \cdots + a_{nn} x_n y_n$$

La expresión formal anterior en las variables x_i, y_i se denomina el *polinomio bilineal* correspondiente a la matriz A. La ecuación [13.1] escrita más adelante muestra que, en cierto sentido, toda forma bilineal es de este tipo.

Denotaremos por $B(V)$ el conjunto de las formas bilineales en V. Se dota a $B(V)$ de una estructura de espacio vectorial definiendo $f + g$ y kf como sigue:

$$(f + g)(u, v) = f(u, v) + g(u, v)$$
$$(kf)(u, v) = kf(u, v)$$

para todos los $f, g \in B(V)$ y $k \in K$. De hecho,

Teorema 13.1: Sea V un espacio vectorial de dimensión n sobre K. Sea $\{\phi_1, \ldots, \phi_n\}$ una base del espacio dual V^*. Entonces $\{f_{ij} : i, j = 1, \ldots, n\}$ es una base de $B(V)$, donde f_{ij} se define según $f_{ij}(u, v) = \phi_i(u)\phi_j(v)$. De este modo, en particular, $\dim B(V) = n^2$.

(Véase la demostración en el Problema 13.4.)

13.3. FORMAS BILINEALES Y MATRICES

Sean f una forma bilineal en V y $S = \{u_1, u_2, \ldots, u_n\}$ una base de V. Supongamos $u, v \in V$ y

$$u = a_1 u_1 + \cdots + a_n u_n \qquad \text{y} \qquad v = b_1 u_1 + \cdots + b_n u_n$$

En tal caso,

$$f(u, v) = f(a_1 u_1 + \cdots + a_n u_n, b_1 u_1 + \cdots + b_n u_n) =$$
$$= a_1 b_1 f(u_1, u_1) + a_1 b_2 f(u_1, u_2) + \cdots + a_n b_n f(u_n, u_n) =$$
$$= \sum_{i,j=1}^{n} a_i b_j f(u_i, u_j)$$

Así pues, f está completamente determinada por los n^2 valores $f(u_i, u_j)$.

La matriz $A = (a_{ij})$, donde $a_{ij} = f(u_i, u_j)$ se llama la *representación matricial de f* relativa a la base S o, simplemente, la *matriz de f* en S. La matriz «representa» a f en el sentido de que

$$f(u, v) = \sum a_i b_j f(u_i, u_j) = (a_1, \ldots, a_n)A\begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix} = [u]_S^T A[v]_S \qquad [13.1]$$

para todos los u, $v \in V$. [Como de costumbre, $[u]_S$ denota el vector (columna) coordenado de $u \in V$ en la base S.]

A continuación nos preguntamos cómo se transforma una matriz que representa a una forma bilineal cuando se selecciona una nueva base. La respuesta se da en el siguiente teorema, demostrado en el Problema 13.6. (Recuérdese que el Teorema 10.4 decía que la matriz de cambio de base P desde una base S hasta otra S' tiene la propiedad de que $[u]_S = P[u]_{S'}$ para todo $u \in V$.)

Teorema 13.2: Sea P la matriz de cambio de base desde una base S hasta otra S'. Si A es la matriz de f en la base original S,

$$B = P^T AP$$

es la matriz de f en la nueva base S'.

El teorema anterior motiva la definición que ahora escribimos.

Definición: Se dice que una matriz B es *congruente* a una matriz A si existe una matriz invertible (o no singular) P tal que $B = P^T AP$.

En virtud del Teorema 13.2, las matrices que representan la misma forma bilineal son, pues, congruentes. Hacemos notar que las matrices congruentes tienen el mismo rango porque P y P^T son no singulares, luego tiene sentido la siguiente definición.

Definición: El *rango* de una forma bilineal f en V, escrito rango f, se define como el rango de cualquiera de sus representaciones matriciales. Decimos que f es *degenerada* o *no degenerada* según sea rango $f <$ dim V o rango $f =$ dim V.

13.4. FORMAS BILINEALES ALTERNADAS

Se dice que una forma bilineal en V es *alternada* si

$$\text{i)} \quad f(v, v) = 0$$

para todo $v \in V$. Si f es alternada,

$$0 = f(u + v, u + v) = f(u, u) + f(u, v) + f v, u) + f(v, v)$$

y así

$$\text{ii)} \quad f(u, v) = -f(v, u)$$

para todos los $u, v \in V$. Una forma bilineal que satisface la condición ii) se dice *antisimétrica* (o *hemisimétrica*). Si $1 + 1 \neq 0$ en K, la condición ii) implica $f(v, v) = -f(v, v)$, lo que a su vez implica la condición i). En otras palabras, las expresiones alternada y antisimétrica son equivalentes cuando $1 + 1 \neq 0$.

El principal teorema de estructura de formas bilineales alternadas, demostrado en el Problema 13.19, es el que se enuncia a continuación.

Teorema 13.3: Sea f una forma bilineal alternada en V. Existe una base de V en la que f se representa por una matriz de la forma

$$\begin{pmatrix} \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} & & & & & & \\ & \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} & & & & & \\ & & \ddots & & & & \\ & & & \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{pmatrix}$$

Más aún, el número de las $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ está unívocamente determinado por f (debido a que es igual a $\frac{1}{2}$ rango f).

En particular, el teorema precedente muestra que una forma bilineal alternada debe tener rango par.

13.5. FORMAS BILINEALES SIMETRICAS. FORMAS CUADRATICAS

Se dice que una forma bilineal f en V es *simétrica* si

$$f(u, v) = f(v, u)$$

para todos los $u, v \in V$. Si A es una representación matricial de f, podemos escribir

$$f(X, Y) = X^T A Y = (X^T A Y)^T = Y^T A^T X$$

(Usamos el hecho de que $X^T A Y$ es un escalar y por tanto es igual a su traspuesta.) Así pues, si f es simétrica,

$$Y^T A^T X = f(X, Y) = f(Y, X) = Y^T A X$$

y como esto es cierto para todos los vectores X, Y, se desprende que $A = A^T$, o sea, que A es simétrica. Recíprocamente, si A es simétrica, f es simétrica.

El principal resultado referente a formas bilineales simétricas, demostrado en el Problema 13.11, se enuncia en el

Teorema 13.4: Sea f una forma bilineal simétrica en V sobre K (en el cual $1 + 1 \neq 0$). Entonces V tiene una base $\{v_1, \ldots, v_n\}$ en la que f se representa por una matriz diagonal, es decir, $f(v_i, v_j) = 0$ para $i \neq j$.

Teorema 13.4 (forma alternativa): Sea A una matriz simétrica sobre K (en el cual $1 + 1 \neq 0$). Entonces existe una matriz invertible (o no singular) P tal que $P^T AP$ es diagonal. Esto es, A es congruente a una matriz diagonal.

Dado que una matriz invertible P es producto de matrices elementales (Teorema 4.10), una manera de obtener la forma diagonal $P^T AP$ es mediante una sucesión de operaciones elementales entre filas y la misma sucesión de operaciones elementales entre columnas. Estas últimas proporcionarán P^T.

Definición: Una aplicación $q: V \to K$ se denomina una forma cuadrática si $q(v) = f(v, v)$ para alguna forma bilineal simétrica f en V.

Ahora bien, si f se representa por una matriz simétrica $A = (a_{ij})$, q se representa como

$$q(X) = f(X, X) = X^T AX = (x_1, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} =$$

$$= \sum_{i,j} a_{ij} x_i x_j = a_{11} x_1^2 + a_{22} x_2^2 + \cdots + a_{nn} x_n^2 + 2 \sum_{i<j} a_{ij} x_i x_j$$

La expresión formal precedente en las variables x_i se denomina el *polinomio cuadrático* correspondiente a la matriz simétrica A. Observemos que si la matriz A es diagonal, q tendrá la *representación diagonal*

$$q(X) = X^T AX = a_{11} x_1^2 + a_{22} x_2^2 + \cdots + a_{nn} x_n^2$$

o sea, el polinomio cuadrático que representa a q no contendrá términos con «productos cruzados». Más aún, según el Teorema 13.4, toda forma cuadrática tiene una representación de esa forma (cuando $1 + 1 \neq 0$).

Si $1 + 1 \neq 0$ en K, la anterior definición puede invertirse, dando $f(u, v) = \frac{1}{2}[q(u + v) - q(u) - q(v)]$, que se conoce como *forma polar* de f.

13.6. FORMAS BILINEALES SIMETRICAS REALES. LEY DE INERCIA

En esta sección trataremos formas bilineales simétricas y formas cuadráticas en espacios vectoriales sobre el cuerpo real $\mathbf{R}$. Estas formas aparecen en muchas ramas de las matemáticas

y la física. La naturaleza especial de **R** permite una teoría independiente. El resultado principal, demostrado en el Problema 13.13, es el que sigue.

Teorema 13.5: Sea f una forma bilineal simétrica en V sobre **R**. Existe una base de V en la que f se representa por una matriz diagonal; cualquier otra representación diagonal tiene el mismo número $\mathbf{p}$ de entradas positivas y el mismo número $\mathbf{n}$ de entradas negativas. La diferencia $\mathbf{s} = \mathbf{p} - \mathbf{n}$ se llama la *signatura* de f.

Una forma bilineal simétrica se dice semidefinida no negativa si

$$q(v) = f(v, v) \geq 0$$

para todo vector v; se dice definida positiva si

$$q(v) = f(v, v) > 0$$

para todo vector $v \neq 0$. Por el Teorema 13.5,

i) f es semidefinida no negativa si y sólo si $\mathbf{s} = \text{rango } (f)$,

ii) f es definida positiva si y sólo si $\mathbf{s} = \dim V$,

siendo $\mathbf{s}$ la signatura de f.

EJEMPLO 13.2. Sea f el producto escalar en $\mathbf{R}^n$; esto es,

$$f(u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

donde $u = (a_i)$ y $v = (b_i)$. Nótese que f es simétrica puesto que

$$f(u, v) = u \cdot v = v \cdot u = f(v, u)$$

Además, f es definida positiva ya que

$$f(u, u) = a_1^2 + a_2^2 + \cdots + a_n^2 > 0$$

cuando $u \neq 0$.

En el Capítulo 14 veremos cómo se transforma una forma cuadrática cuando la matriz de transición P es ortogonal. Si P es tan sólo no singular, q puede representarse en forma diagonal con únicamente 1 y -1 como coeficientes no nulos. Concretamente,

Corolario 13.6: Cualquier forma cuadrática real tiene una única representación de la forma

$$q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_{\mathbf{p}}^2 - x_{\mathbf{p}+1}^2 - \cdots - x_r^2$$

donde $r = \mathbf{p} + \mathbf{n}$ es el rango de la forma.

A veces se llama ley de inercia o teorema de Sylvester al resultado anterior para formas cuadráticas reales.

13.7. FORMAS HERMITICAS

Sea V un espacio vectorial de dimensión finita sobre el cuerpo complejo **C**. Sea $f: V \times V \to \mathbf{C}$ tal que

i) $f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$
ii) $f(u, v) = \overline{f(v, u)}$

donde $a, b \in \mathbf{C}$ y $u_i, v \in V$. Entonces f se denomina una *forma hermítica* en V. (Como de costumbre, $\bar{k}$ denota el complejo conjugado de $k \in \mathbf{C}$.) Por i) y ii),

$$f(u, av_1 + bv_2) = \overline{f(av_1 + bv_2, u)} = \overline{af(v_1, u) + bf(v_2, u)} =$$
$$= \bar{a}\overline{f(v_1, u)} + \bar{b}\overline{f(v_2, u)} = \bar{a}f(u, v_1) + \bar{b}f(u, v_2)$$

Esto es,

iii) $f(u, av_1 + bv_2) = \bar{a}f(u, v_1) + \bar{b}f(u, v_2)$

Como antes, expresamos la condición i) diciendo que f es lineal en la primera variable. Por otra parte, expresamos la condición iii) diciendo que f es *antilineal* en la segunda variable. Nótese que según ii) tenemos $f(v, v) = \overline{f(v, v)}$ y así $f(v, v)$ es real para todo $v \in V$.

Los resultados de las Secciones 13.5 y 13.6 para formas simétricas tienen sus análogos para formas hermíticas. Así la aplicación $q: V \to \mathbf{R}$, definida por $q(v) = f(v, v)$, se conoce como la *forma cuadrática hermítica* o *forma cuadrática compleja* asociada a la forma hermítica f. Podemos obtener f a partir de q en la forma polar:

$$f(u, v) = \tfrac{1}{4}[q(u + v) - q(u - v)] + \tfrac{1}{4}[q(u + iv) - q(u - iv)]$$

Supongamos ahora que $S = \{u_1, \ldots, u_n\}$ es una base de V. La matriz $H = (h_{ij})$ con $h_{ij} = f(u_i, u_j)$ se denomina la *representación matricial* de f en la base S. Por ii), $f(u_i, u_j) = \overline{f(u_j, u_i)}$, luego H es hermítica y en particular sus entradas diagonales son reales. De este modo, toda representación diagonal de f contiene sólo entradas reales. El próximo teorema, que se demostrará en el Problema 13.33, es el análogo complejo del Teorema 13.5 referente a formas bilineales reales simétricas.

Teorema 13.7: Sea f una forma hermítica en V. Existe una base $S = \{u_1, \ldots, u_n\}$ de V en la que f se representa por una matriz diagonal, es decir, $f(u_i, u_j) = 0$ para $i \neq j$. Más aún, toda representación diagonal de f tiene el mismo número **p** de entradas positivas y el mismo número **n** de entradas negativas. La diferencia $\mathbf{s} = \mathbf{p} - \mathbf{n}$ se llama la signatura de f.

Análogamente, una forma hermítica f se dice semidefinida no negativa si

$$q(v) = f(v, v) \geq 0$$

para todo $v \in V$, y se dice definida positiva si

$$q(v) = f(v, v) > 0$$

para todo $v \neq 0$.

EJEMPLO 13.3. Sea f el producto escalar en $\mathbf{C}^n$; esto es,

$$f(u, v) = u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

donde $u = (z_i)$ y $v = (w_i)$. Entonces f es una forma hermítica en $\mathbf{C}^n$. Además, f es definida positiva porque, para todo $v \neq 0$,

$$f(u, u) = z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 > 0$$

PROBLEMAS RESUELTOS

FORMAS BILINEALES

13.1. Sean $u = (x_1, x_2, x_3)$ y $v = (y_1, y_2, y_3)$, y sea

$$f(u, v) = 3x_1y_1 - 2x_1y_2 + 5x_2y_1 + 7x_2y_2 - 8x_2y_3 + 4x_3y_2 - x_3y_3$$

Expresar f en notación matricial.

Sea A la matriz 3×3 cuya entrada ij es el coeficiente de x_iy_j. En ese caso,

$$f(u, v) = X^TAY = (x_1, x_2, x_3)\begin{pmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

13.2. Sea A una matriz $n \times n$ sobre K. Mostrar que la aplicación f definida por

$$f(X, Y) = X^TAY$$

es una forma bilineal en K^n.

Para todos los $a, b \in K$ y todos los $X_i, Y_i \in K^n$,

$$f(aX_1 + bX_2, Y) = (aX_1 + bX_2)^TAY = (aX_1^T + bX_2^T)AY =$$
$$= aX_1^TAY + bX_2^TAY = af(X_1, Y) + bf(X_2, Y)$$

Por tanto, f es lineal en la primera variable. Asimismo,

$$f(X, aY_1 + bY_2) = X^TA(aY_1 + bY_2) = aX^TAY_1 + bX^TAY_2 = af(X, Y_1) + bf(X, Y_2)$$

Por tanto, f es lineal en la segunda variable y es pues una forma bilineal en K^n.

13.3. Sea f la forma bilineal en $\mathbf{R}^2$ definida por

$$f((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$$

a) Hallar la matriz A de f en la base $\{u_1 = (1, 0),\ u_2 = (1, 1)\}$.

b) Hallar la matriz B de f en la base $\{v_1 = (2, 1),\ v_2 = (1, -1)\}$.

c) Encontrar la matriz de cambio de base P desde la base $\{u_i\}$ hasta la base $\{v_i\}$ y verificar que $B = P^T AP$.

a) Tomamos $A = (a_{ij})$, donde $a_{ij} = f(u_i, u_j)$:

$$\begin{aligned} a_{11} &= f(u_1, u_1) = f((1, 0), (1, 0)) = 2 - 0 + 0 = 2 \\ a_{12} &= f(u_1, u_2) = f((1, 0), (1, 1)) = 2 - 3 + 0 = -1 \\ a_{21} &= f(u_2, u_1) = f((1, 1), (1, 0)) = 2 - 0 + 0 = 2 \\ a_{22} &= f(u_2, u_2) = f((1, 1), (1, 1)) = 2 - 3 + 1 = 0 \end{aligned}$$

Así $A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$ es la matriz de f en la base $\{u_1, u_2\}$.

b) Tomamos $B = (b_{ij})$, donde $b_{ij} = f(v_i, v_j)$:

$$\begin{aligned} b_{11} &= f(v_1, v_1) = f((2, 1), (2, 1)) = 8 - 6 + 1 = 3 \\ b_{12} &= f(v_1, v_2) = f((2, 1), (1, -1)) = 4 + 6 - 1 = 9 \\ b_{21} &= f(v_2, v_1) = f((1, -1), (2, 1)) = 4 - 3 - 1 = 0 \\ b_{22} &= f(v_2, v_2) = f((1, -1), (1, -1)) = 2 + 3 + 1 = 6 \end{aligned}$$

Así $B = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$ es la matriz de f en la base $\{v_1, v_2\}$.

c) Debemos escribir v_1 y v_2 en términos de los u_i:

$$\begin{aligned} v_1 &= (2, 1) = (1, 0) + (1, 1) = u_1 + u_2 \\ v_2 &= (1, -1) = 2(1, 0) - (1, 1) = 2u_1 - u_2 \end{aligned}$$

Entonces $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$, por lo que $P^T = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$. De este modo

$$P^T AP = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix} = B$$

13.4. Demostrar el Teorema 13.1.

Sea $\{u_1, \ldots, u_n\}$ la base de V dual a $\{\phi_i\}$. Mostremos primero que $\{f_{ij}\}$ genera $B(V)$. Sea $f \in B(V)$ y supongamos $f(u_i, u_j) = a_{ij}$. Afirmamos que $f = \sum a_{ij} f_{ij}$. Basta probar que

$$f(u_s, u_t) = (\textstyle\sum a_{ij} f_{ij})(u_s, u_t) \qquad \text{para} \quad s, t = 1, \ldots, n$$

Tenemos

$$(\textstyle\sum a_{ij} f_{ij})(u_s, u_t) = \sum a_{ij} f_{ij}(u_s, u_t) = \sum a_{ij} \phi_i(u_s)\phi_j(u_t) = \sum a_{ij} \delta_{is} \delta_{jt} = a_{st} = f(u_s, u_t)$$

como se pedía. Por consiguiente, $\{f_{ij}\}$ genera $B(V)$. A continuación supongamos $\sum a_{ij} f_{ij} = 0$. En ese caso, para $s, t = 1, \ldots, n$,

$$0 = 0(u_s, u_t) = (\textstyle\sum a_{ij} f_{ij})(u_s, u_t) = a_{rs}$$

El último paso se obtiene como antes. Así pues, $\{f_{ij}\}$ es independiente y por ende es una base de $B(V)$.

13.5. Denótese por $[f]$ la representación matricial de una forma bilineal f en V relativa a una base $\{u_1, \ldots, u_n\}$. Mostrar que la aplicación $f \mapsto [f]$ es un isomorfismo de $B(V)$ sobre el espacio vectorial las matrices n-cuadradas

Dado que f está completamente determinada por los escalares $f(u_i, u_j)$, la aplicación $f \mapsto [f]$ es inyectiva y suprayectiva. Basta probar que la aplicación $f \mapsto [f]$ es un homomorfismo; esto es, que

$$[af + bg] = a[f] + b[g] \qquad (*)$$

No obstante, para $i, j = 1, \ldots, n$,

$$(af + bg)(u_i, u_j) = af(u_i, u_j) + bg(u_i, u_j)$$

que es otra forma de escribir $(*)$. Queda así demostrado el resultado.

13.6. Demostrar el Teorema 13.2.

Sean $u, v \in V$. Siendo P la matriz de cambio de base de S a S' tenemos $P[u]_{S'} = [u]_S$ y también $P[v]_{S'} = [v]_S$, luego $[u]_S^T = [u]_{S'}^T P^T$. De este modo,

$$f(u, v) = [u]_S^T A[v]_S = [u]_{S'}^T P^T AP[v]_{S'}$$

Como u y v son elementos arbitrarios en V, $P^T AP$ es la matriz de f en la base S'.

FORMAS BILINEALES SIMETRICAS. FORMAS CUADRATICAS

13.7. Hallar la matriz simétrica correspondiente a cada uno de los siguientes polinomios cuadráticos:

a) $q(x, y, z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$ *b)* $q(x, y, z) = x^2 - 2yz + xz$.

La matriz simétrica $A = (a_{ij})$ que representa a $q(x_1, \ldots, x_n)$ tiene la entrada diagonal a_{ii} igual al coeficiente de x_i^2, y las entradas a_{ij} y a_{ji} iguales cada una a la mitad del coeficiente de $x_i x_j$. Así pues,

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & -1 & -3 \\ 4 & -3 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 0 & -1 \\ \frac{1}{2} & -1 & 0 \end{pmatrix}$$

a) *b)*

13.8. Para la matriz real simétrica A escrita a continuación, encontrar una matriz no singular P tal que $P^T AP$ sea diagonal y hallar su signatura.

$$A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$$

Comenzamos por construir la matriz por bloques (A, I):

$$(A, I) = \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right)$$

Efectuamos las operaciones entre filas $3R_1 + R_2 \to R_2$ y $-2R_1 + R_3 \to R_3$ sobre (A, I) y después las correspondientes operaciones entre columnas $3C_1 + C_2 \to C_2$ y $-2C_1 + C_3 \to C_3$ sobre A obteniendo

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

Acto seguido, efectuamos la operación entre filas $R_2 + 2R_3 \to R_3$ y después la correspondiente operación entre columnas $C_2 + 2C_3 \to C_3$ para llegar a

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \quad \text{y luego a} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

Se ha diagonalizado A. Tomemos $P = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$; entonces $P^T AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix}$.

La signatura de A es $s = 2 - 1 = 1$.

13.9. Supóngase $1 + 1 \neq 0$ en K. Dar un algoritmo formal para diagonalizar (bajo congruencia) una matriz simétrica $A = (a_{ij})$ sobre K.

Caso i): $a_{11} \neq 0$. Efectuamos las operaciones entre filas $-a_{i1}R_1 + a_{11}R_i \to R_i$, $i = 2, \ldots, n$, seguidas de las correspondientes operaciones entre columnas $-a_{i1}C_1 + a_{11}C_i \to C_i$ para reducir A a la forma $\begin{pmatrix} a_{11} & 0 \\ 0 & B \end{pmatrix}$.

Caso ii): $a_{11} = 0$ pero $a_{ii} \neq 0$ para algún $i > 1$. Efectuamos la operación entre filas $R_1 \leftrightarrow R_i$ y después la correspondiente entre columnas $C_1 \leftrightarrow C_i$ para llevar a_{ii} a la primera posición diagonal. Esto reduce la matriz a la de i).

Caso iii): Todas las entradas diagonales $a_{ii} = 0$. Elegimos i, j tales que $a_{ij} \neq 0$ y efectuamos la operación entre filas $R_j + R_i \to R_i$ seguida de la correspondiente operación entre columnas $C_j + C_i \to C_i$ para llevar $2a_{ij} \neq 0$ a la i-ésima posición diagonal. Esto reduce la matriz a la de ii).

En cada uno de los casos podemos reducir finalmente A a la forma $\begin{pmatrix} a_{11} & 0 \\ 0 & B \end{pmatrix}$, donde B es una matriz simétrica de orden menor que el de A. Por inducción podremos acabar llevando A a forma diagonal.

Nota: La hipótesis de que $1 + 1 \neq 0$ en K se ha usado en iii), donde afirmamos que $2a_{ij} \neq 0$.

13.10. Sea q la forma cuadrática asociada a la forma bilineal simétrica f. Comprobar la identidad polar $f(u, v) = \frac{1}{2}[q(u + v) - q(u) - q(v)]$. (Supóngase que $1 + 1 \neq 0$.)

Tenemos

$$\begin{aligned} q(u + v) - q(u) - q(v) &= f(u + v, u + v) - f(u, u) - f(v, v) = \\ &= f(u, u) + f(u, v) + f(v, u) + f(v, v) - f(u, u) - f(v, v) = \\ &= 2f(u, v) \end{aligned}$$

Si $1 + 1 \neq 0$, podemos dividir por 2 para conseguir la identidad requerida.

13.11. Demostrar el Teorema 13.4.

Método 1. Si $f = 0$ o si $\dim V = 1$, el teorema se cumple claramente. Por tanto, podemos suponer $f \neq 0$ y $\dim V = n > 1$. Si $q(v) = f(v, v) = 0$ para todo $v \in V$, la forma polar de f (véase el Problema 13.10) implica $f = 0$. Por esta razón, podemos suponer que existe un vector $v_1 \in V$ tal que $f(v_1, v_1) \neq 0$. Sean U el subespacio generado por v_1 y W el conjunto de aquellos vectores $v \in V$ para los que $f(v_1, v) = 0$. Afirmamos que $V = U \oplus W$.

i) Demostración de que $U \cap W = \{0\}$: Supongamos $u \in U \cap W$. Como $u \in U$, $u = kv_1$ para algún escalar $k \in K$. Como $u \in W$, $0 = f(u, u) = f(kv_1, kv_1) = k^2 f(v_1, v_1)$. Pero $f(v_1, v_1) \neq 0$, luego $k = 0$ y por consiguiente $u = kv_1 = 0$. De este modo, $U \cap W = \{0\}$.

ii) Demostración de que $V = U + W$: sea $v \in V$. Tomamos

$$w = v - \frac{f(v_1, v)}{f(v_1, v_1)} v_1 \qquad [1]$$

Entonces

$$f(v_1, w) = f(v_1, v) - \frac{f(v_1, v)}{f(v_1, v_1)} f(v_1, v_1) = 0$$

Siendo así, $w \in W$. Por [1], v es la suma de un elemento de U y uno de W, de manera que $V = U + W$. Por i) y ii), $V = U \oplus W$.

Ahora f restringida a W es una forma bilineal simétrica en W. Pero $\dim W = n - 1$, luego, por inducción, existe una base $\{v_2, \ldots, v_n\}$ de W tal que $f(v_i, v_j) = 0$ para $i \neq j$ y $2 \leq i, j \leq n$. Por la propia definición de W, $f(v_1, v_j) = 0$ para $j = 2, \ldots, n$. Por tanto, la base $\{v_1, \ldots, v_n\}$ de V tiene la propiedad requerida de que $f(v_i, v_j) = 0$ para $i \neq j$.

Método 2. El algoritmo del Problema 13.9 muestra que toda matriz simétrica sobre K es congruente a una matriz diagonal. Esto es equivalente a la afirmación de que f tiene una representación diagonal.

13.12. Sea $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \cdots & \\ & & & a_n \end{pmatrix}$ una matriz diagonal sobre K. Probar que:

a) Para escalares no nulos cualesquiera $k_1, \ldots, k_n \in K$, A es congruente a una matriz diagonal con entradas diagonales $a_i k_i^2$.

b) Si K es el cuerpo complejo **C**, A es congruente a una matriz diagonal con sólo 1 y 0 en las entradas diagonales.

c) Si K es el cuerpo real **R**, A es congruente a una matriz diagonal con sólo 1, -1 y 0 en las entradas diagonales.

a) Sea P la matriz diagonal con entradas diagonales k_i. Entonces

$$P^T A P = \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \cdots & \\ & & & k_n \end{pmatrix} \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \cdots & \\ & & & a_n \end{pmatrix} \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \cdots & \\ & & & k_n \end{pmatrix} = \begin{pmatrix} a_1 k_1^2 & & & \\ & a_2 k_2^2 & & \\ & & \cdots & \\ & & & a_n k_n^2 \end{pmatrix}$$

b) Sea P la matriz diagonal con entradas diagonales $b_i = \begin{cases} 1/\sqrt{a_i} & \text{si } a_i \neq 0 \\ 1 & \text{si } a_i = 0 \end{cases}$. En tal caso, $P^T AP$ tiene la forma pedida.

c) Sea P la matriz diagonal con entradas diagonales $b_i = \begin{cases} 1/\sqrt{|a_i|} & \text{si } a_i \neq 0 \\ 1 & \text{si } a_i = 0 \end{cases}$. En tal caso, $P^T AP$ tiene la forma pedida.

Nota: Subrayamos el hecho de que *b*) deja de ser cierto si se reemplaza congruencia por congruencia hermítica (véanse los Problemas 13.32 y 13.33).

13.13. Demostrar el Teorema 13.5.

En virtud del Teorema 13.4, existe una base $\{u_1, \ldots, u_n\}$ de V en la que f se representa por una matriz diagonal con, digamos, $\mathbf{p}$ entradas positivas y $\mathbf{n}$ negativas. Supongamos ahora que $\{w_1, \ldots, w_n\}$ es otra base de V en la que f se representa por una matriz diagonal con, digamos, $\mathbf{p}'$ entradas positivas y $\mathbf{n}'$ negativas. Podemos suponer, sin pérdida de generalidad, que en cada matriz aparecen primero las entradas positivas. Puesto que rango $f = \mathbf{p} + \mathbf{n} = \mathbf{p}' + \mathbf{n}'$, basta probar que $\mathbf{p} = \mathbf{p}'$

Sean U la envolvente lineal de $u_1, \ldots, u_{\mathbf{p}}$ y W la de $w_{\mathbf{p}'+1}, \ldots, w_n$. Entonces $f(v, v) > 0$ para todo $v \in U$ no nulo y $f(v, v) \leq 0$ para todo $v \in W$ no nulo. De aquí $U \cap W = \{0\}$. Nótese que $\dim U = \mathbf{p}$ y $\dim W = n - \mathbf{p}'$. Siendo así,

$$\dim (U + W) = \dim U + \dim W - \dim (U \cap W) = \mathbf{p} + (n - \mathbf{p}') - 0 = \mathbf{p} - \mathbf{p}' + n$$

Pero $\dim (U + W) \leq \dim V = n$, luego $\mathbf{p} - \mathbf{p}' + n \leq n$ o $\mathbf{p} \leq \mathbf{p}'$. De forma similar, $\mathbf{p}' \leq \mathbf{p}$ y por consiguiente $\mathbf{p} = \mathbf{p}'$, como se pedía.

Nota: El teorema y la demostración precedentes sólo dependen del concepto de positividad. El teorema será cierto, pues, para cualquier subcuerpo K del cuerpo real $\mathbf{R}$, como es el cuerpo racional $\mathbf{Q}$.

13.14. Se dice que una matriz real simétrica $n \times n$ es *definida positiva* si $X^T AX > 0$ para todo vector (columna) no nulo $X \in \mathbf{R}^n$, es decir, si A es definida positiva vista como forma bilineal. Sea B cualquier matriz real no singular. Mostrar que: *a*) $B^T B$ es simétrica, *b*) $B^T B$ es definida positiva.

a) $(B^T B)^T = B^T B^{TT} = B^T B$, luego $B^T B$ es simétrica.

b) Dado que B es no singular, $BX \neq 0$ para todo $X \in \mathbf{R}^n$ no nulo. Por tanto, el producto escalar de BX consigo mismo, $BX \cdot BX = (BX)^T (BX)$, es positivo. De este modo, $X^T (B^T B) X = (X^T B^T) = (BX)^T (BX) > 0$, como se pedía.

FORMAS HERMITICAS

13.15. Determinar cuáles de las siguientes matrices son hermíticas:

$$\begin{pmatrix} 2 & 2+3i & 4-5i \\ 2-3i & 5 & 6+2i \\ 4+5i & 6-2i & -7 \end{pmatrix} \qquad \begin{pmatrix} 3 & 2-i & 4+i \\ 2-i & 6 & i \\ 4+i & i & 3 \end{pmatrix} \qquad \begin{pmatrix} 4 & -3 & 5 \\ -3 & 2 & 1 \\ 5 & 1 & -6 \end{pmatrix}$$

a) *b*) *c*)

Una matriz $A = (a_{ij})$ es hermítica si y sólo si $A = A^*$, o sea, si y sólo si $a_{ij} = \overline{a_{ji}}$.

a) La matriz es hermítica, ya que es igual a su traspuesta conjugada.

b) La matriz no es hermítica, incluso a pesar de ser simétrica.

c) La matriz es hermítica. De hecho, una matriz real es hermítica si y sólo si es simétrica.

13.16. Sea A una matriz hermítica. Probar que f es una forma hermítica en $\mathbf{C}^n$, donde f se define por $f(X, Y) = X^T A \bar{Y}$.

Para todos los $a, b \in \mathbf{C}$ y todos los $X_1, X_2, Y \in \mathbf{C}^n$,

$$f(aX_1 + bX_2, Y) = (aX_1 + bX_2)^T A\bar{Y} = (aX_1^T + bX_2^T)A\bar{Y} =$$
$$= aX_1^T A\bar{Y} + bX_2^T A\bar{Y} = af(X_1, Y) + bf(X_2, Y)$$

Por consiguiente, f es lineal en la primera variable. Asimismo,

$$\overline{f(X, Y)} = \overline{X^T A \bar{Y}} = \overline{(X^T A \bar{Y})^T} = \overline{\bar{Y}^T A^T X} = Y^T A^* \bar{X} = Y^T A \bar{X} = f(Y, X)$$

De aquí que f sea una forma hermítica en $\mathbf{C}^n$. (*Nota:* Usamos el hecho de que $X^T A \bar{Y}$ es un escalar, por lo que es igual a su traspuesta.)

13.17. Sea f una forma hermítica en V. Sea H la matriz de f en una base $S = \{u_1, \ldots, u_n\}$ de V. Probar que:

a) $f(u, v) = [u]_S^T H \overline{[v]_S}$ para todos los $u, v \in V$.

b) Si P es la matriz de cambio de base desde S hasta una nueva base S' de V, $B = P^T H \bar{P}$ (o $B = Q^* HQ$, donde $Q = \bar{P}$) es la matriz de f en la nueva base S'.

Nótese que *b*) es el análogo complejo del Teorema 13.2.

a) Sean $u, v \in V$ y supongamos $u = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$ y $v = b_1 u_1 + b_2 u_2 + \cdots + b_n u_n$. En tal caso,

$$f(u, v) = f(a_1 u_1 + \cdots + a_n u_n, b_1 u_1 + \cdots + b_n u_n) =$$

$$= \sum_{i,j} a_i \bar{b}_j f(u_i u_j) = (a_1, \ldots, a_n) H \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \cdots \\ \bar{b}_n \end{pmatrix} = [u]_S^T H \overline{[v]_S}$$

como se pedía.

b) Al ser P la matriz de cambio de base desde S hasta S', tenemos $P[u]_{S'} = [u]_S$ y $P[v]_{S'} = [v]_S$; por tanto, $[u]_S^T = [u]_{S'}^T P^T$ y $\overline{[v]_S} = \overline{P[v]_{S'}}$. Así, por *a*),

$$f(u, v) = [u]_S^T H \overline{[v]_S} = [u]_{S'}^T P^T H \bar{P} \overline{[v]_{S'}}$$

Pero u y v son elementos arbitrarios de V, luego $P^T H \bar{P}$ es la matriz de f en la base S'.

13.18. Sea $H = \begin{pmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{pmatrix}$ una matriz hermítica. Hallar una matriz no singular P tal que $P^T H \bar{P}$ sea diagonal.

Primero construimos la matriz por bloques (H, I):

$$\left(\begin{array}{ccc|ccc} 1 & 1+i & 2i & 1 & 0 & 0 \\ 1-i & 4 & 2-3i & 0 & 1 & 0 \\ -2i & 2+3i & 7 & 0 & 0 & 1 \end{array}\right)$$

Efectuamos las operaciones entre filas $(-1+i)R_1 + R_2 \to R_2$ y $2iR_1 + R_3 \to R_3$ sobre (H, I) seguidas de las correspondientes «operaciones entre columnas conjugadas» $(-1-i)C_1 + C_2 \to C_2$ y $-2iC_1 + C_3 \to C_3$ sobre H obteniendo

$$\left(\begin{array}{ccc|ccc} 1 & 1+i & 2i & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{array}\right)$$

A continuación efectuamos la operación entre filas $R_3 \to -5iR_2 + 2R_3$ y la conjugada $C_3 \to 5iC_2 + 2C_3$ obteniendo

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 0 & -19 & 5+9i & -5i & 2 \end{array}\right) \quad \text{y luego} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1+i & 1 & 0 \\ 0 & 0 & -38 & 5+9i & -5i & 2 \end{array}\right)$$

Se ha diagonalizado H. Tomemos

$$P = \begin{pmatrix} 1 & -1+i & 5+9i \\ 0 & 1 & -5i \\ 0 & 0 & 2 \end{pmatrix} \quad \text{y entonces} \quad P^T H \bar{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$$

Adviértase que la signatura $\mathbf{s}$ de H es $\mathbf{s} = 2 - 1 = 1$.

PROBLEMAS VARIOS

13.19. Demostrar el Teorema 13.3.

Si $f = 0$, el teorema es obviamente cierto. Además, si $\dim V = 1$, $f(k_1 u, k_2 u) = k_1 k_2 f(u, u) = 0$ y así $f = 0$. En consecuencia, podemos suponer que $\dim V > 1$ y $f \neq 0$.

Puesto que $f \neq 0$, existen $u_1, u_2 \in V$ (no nulos) tales que $f(u_1, u_2) \neq 0$. De hecho, multiplicando u_1 por un factor apropiado, podemos suponer que $f(u_1, u_2) = 1$, luego $f(u_2, u_1) = -1$. Ahora bien, u_1 y u_2 son linealmente independientes, porque si, digamos, $u_2 = ku_1$, $f(u_1, u_2) = f(u_1, ku_1) = kf(u_1, u_1) = 0$. Sea U el subespacio generado por u_1 y u_2. Nótese que:

i) La representación matricial de la restricción de f a U en la base $\{u_1, u_2\}$ es $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

ii) Si $u \in U$, digamos $u = au_1 + bu_2$,

$$f(u, u_1) = f(au_1 + bu_2, u_1) = -b$$
$$f(u, u_2) = f(au_1 + bu_2, u_2) = a$$

Sea W el conjunto formado por aquellos vectores $w \in W$ tales que $f(w, u_1) = 0$ y $f(w, u_2) = 0$ Equivalentemente,

$$W = \{w \in V : f(w, u) = 0 \text{ para todo } u \in U\}$$

Afirmamos que $V = U \oplus W$. Es claro que $U \cap W = \{0\}$, por lo que solamente resta probar que $V = U + W$. Sea $v \in V$. Tomemos

$$u = f(v, u_2)u_1 - f(v, u_1)u_2 \qquad \text{y} \qquad w = v - u$$

Dado que u es combinación lineal de u_1 y u_2, $u \in U$. Probemos que $w \in W$. Por [1] y ii), $f(u, u_1) = f(v, u_1)$, luego

$$f(w, u_1) = f(v - u, u_1) = f(v, u_1) - f(u, u_1) = 0$$

Similarmente, $f(u, u_2) = f(v, u_2)$, de modo que

$$f(w, u_2) + f(v - u, u_2) = f(v, u_2) - f(u, u_2) = 0$$

Entonces $w \in W$, y así, por [1], $v = u + w$ con $u \in U$ y $w \in W$. Esto muestra que $V = U + W$ y por tanto $V = U \oplus W$.

Ahora la restricción de f a W es una forma bilineal alternada en W. Por inducción, existe una base $u_3, \ldots, u_n$ de W en la que la matriz que representa a f restringida a W tiene la forma deseada. De acuerdo con esto, $u_1, u_2, u_3, \ldots, u_n$ es una base de V en la que la matriz que representa a f tiene la forma deseada.

PROBLEMAS SUPLEMENTARIOS

FORMAS BILINEALES

13.20. Sea V el espacio vectorial de las matrices 2×2 sobre $\mathbf{R}$. Sean $M = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ y $f(A, B) = \operatorname{tr} A^T MB$, donde $A, B \in V$ y «tr» denota la traza. *a*) Mostrar que f es una forma bilineal en V. *b*) Hallar la matriz de f en la base $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$.

13.21. Sea $B(V)$ el conjunto de las formas bilineales en V sobre K. Demostrar:

a) Si $f, g \in B(V)$, $f + g$ y kf, con $k \in K$, también pertenecen a $B(V)$. Así que $B(V)$ es un subespacio del espacio vectorial de todas las funciones de $V \times V$ en K.

b) Si ϕ y σ son funcionales lineales en V, $f(u, v) = \phi(u)\sigma(v)$ pertenece a $B(V)$.

13.22. Sea f una forma bilineal en V. Para cualquier subconjunto S de V escribimos

$$S^\perp = \{v \in V : f(u, v) = 0 \text{ para todo } u \in S\} \qquad S^\top = \{v \in V : f(v, u) = 0 \text{ para todo } u \in S\}$$

Probar que: *a*) $S^\perp$ y $S^\top$ son subespacios de V; *b*) $S_1 \subseteq S_2$ implica $S_2^\perp \subseteq S_1^\perp$ y $S_2^\top \subseteq S_1^\top$; *c*) $\{0\}^\perp = \{0\}^\top = V$.

13.23. Demostrar la siguiente aserción: si f es una forma bilineal en V, rango $f = \dim V - \dim V^\perp = \dim V - \dim V^\top$ y por tanto $\dim V^\perp = \dim V^\top$.

13.24. Sea f una forma bilineal en V. Para cada $u \in V$, defínanse $\hat{u}: V \to K$ y $\tilde{u}: V \to K$ según $\hat{u}(x) = f(x, u)$ y $\tilde{u}(x) = f(u, x)$. Demostrar:

a) $\hat{u}$ y $\tilde{u}$ son lineales, es decir, $\hat{u}, \tilde{u} \in V^*$.

b) $u \mapsto \hat{u}$ y $u \mapsto \tilde{u}$ son aplicaciones lineales de V en V^*.

c) rango f = rango $(u \mapsto \hat{u})$ = rango $(u \mapsto \tilde{u})$.

13.25. Mostrar que la congruencia de matrices es una relación de equivalencia, o sea: i) A es congruente a A; ii) si A es congruente a B, B es congruente a A; iii) si A es congruente a B y B es congruente a C, A es congruente a C.

FORMAS BILINEALES SIMETRICAS. FORMAS CUADRATICAS

13.26. Hallar la matriz simétrica asociada a cada uno de los polinomios cuadráticos:

a) $q(x, y, z) = 2x^2 - 8xy + y^2 - 16xz + 14yz + 5z^2$

b) $q(x, y, z) = x^2 - xz + y^2$

c) $q(x, y, z) = xy + y^2 + 4xz + z^2$

d) $q(x, y, z) = xy + yz$

13.27. Para cada una de las siguientes matrices A, encontrar una matriz no singular P tal que $P^T AP$ sea diagonal:

a) $A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$, *b*) $A = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 6 & -9 \\ 3 & -9 & 4 \end{pmatrix}$, *c*) $A = \begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -5 & -1 \\ -2 & -5 & 6 & 9 \\ -3 & -1 & 9 & 11 \end{pmatrix}$

En cada caso, hallar el rango y la signatura.

13.28. Sea $S(V)$ el conjunto de las formas bilineales simétricas en V. Probar que:

i) $S(V)$ es un subespacio de $B(V)$; ii) si dim $V = n$, dim $S(V) = \frac{1}{2}n(n + 1)$.

13.29. Supóngase que A es una matriz real simétrica definida positiva. Mostrar que existe una matriz no singular P tal que $A = P^T P$.

13.30. Considérese un polinomio real cuadrático $q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_i x_j$, donde $a_{ij} = a_{ji}$.

i) Si $a_{11} \neq 0$, probar que la sustitución

$$x_1 = y_1 - \frac{1}{a_{11}}(a_{12}y_2 + \cdots + a_{1n}y_n), \; x_2 = y_2, \ldots, x_n = y_n$$

conduce a la ecuación $q(x_1, \ldots, x_n) = a_{11}y_1^2 + q'(y_2, \ldots, y_n)$, donde q' es también un polinomio cuadrático.

ii) Si $a_{11} = 0$ pero, por ejemplo, $a_{12} \neq 0$, probar que la sustitución

$$x_1 = y_1 + y_2, \; x_2 = y_1 - y_2, \; x_3 = y_3, \ldots, x_n = y_n$$

conduce a la ecuación $q(x_1, \ldots, x_n) = \sum b_{ij}y_i y_j$, donde $b_{11} \neq 0$, esto es, reduce este caso al i). Este método de diagonalización de q se conoce como «completar el cuadrado».

FORMAS HERMITICAS

13.31. Sea A cualquier matriz compleja no singular. Mostrar que $H = A^* A$ es hermítica y definida positiva.

13.32. Decimos que B es *congruente hermítica* a A si existe una matriz no singular Q tal que $B = Q^* AQ$. Probar que la congruencia hermítica es una relación de equivalencia.

13.33. Demostrar el Teorema 13.7. [Nótese que la segunda parte del teorema no se cumple para formas bilineales simétricas complejas, como muestra el Problema 13.12 ii). No obstante, la demostración del Teorema 13.5 en el Problema 13.13 puede trasladarse al caso hermítico.]

PROBLEMAS VARIOS

13.34. Sean V y W espacios vectoriales sobre K. Una aplicación $f: V \times W \to K$ se llama una *forma bilineal* en V y W si:

i) $f(av_1 + bv_2, w) = af(v_1, w) + bf(v_2, w)$

ii) $f(v, aw_1 + bw_2) = af(v, w_1) + bf(v, w_2)$

para todos los $a, b \in K$, $v_i \in V$, $w_j \in W$. Demostrar lo siguiente:

a) El conjunto $B(V, W)$ de las formas bilineales en V y W es un subespacio del espacio vectorial de las funciones de $V \times W$ en K.

b) Si $\{\phi_1, \ldots, \phi_m\}$ es una base de V^* y $\{\sigma_1, \ldots, \sigma_n\}$ una de W^*, $\{f_{ij}: i = 1, \ldots, m, j = 1, \ldots, n\}$ es una base de $B(V, W)$, donde f_{ij} se define por $f_{ij}(v, w) = \phi_i(v)\sigma_j(w)$. Siendo así, $\dim B(V, W) = \dim V \cdot \dim W$.

Nota: Obsérvese que si $V = W$, obtenemos el espacio $B(V)$ estudiado en este capítulo.

13.35. Sea V un espacio vectorial sobre K. Una aplicación $f: \overbrace{V \times V \times \cdots \times V}^{m \text{ veces}} \to K$ se denomina una *forma multilineal* (o *m-lineal*) en V si f es lineal en cada variable, es decir, si para $i = 1, \ldots, m$,

$$f(\ldots, \widehat{au + bv}, \ldots) = af(\ldots, \hat{u}, \ldots) + bf(\ldots, \hat{v}, \ldots)$$

donde $\hat{}$ denota la variable i-ésima, manteniéndose fijas el resto de las variables. Se dice que una forma m-lineal f es *alternada* si

$$f(v_1, \ldots, v_m) = 0 \qquad \text{siempre que} \qquad v_i = v_k,\ i \neq k$$

Demostrar:

a) El conjunto $B_m(V)$ de las formas m-lineales en V es un subespacio del espacio vectorial de las funciones de $V \times V \times \cdots \times V$ en K.

b) El conjunto $A_m(V)$ de las formas m-lineales alternadas en V es un subespacio de $B_m(V)$.

Nota 1: Si $m = 2$, obtenemos el espacio $B(V)$ estudiado en este capítulo.

Nota 2: Si $V = K^m$, la función determinante es una forma m-lineal alternada particular en V.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

13.20. *b*) $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$

13.26. *a*) $\begin{pmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{pmatrix}$, *b*) $\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}$, *c*) $\begin{pmatrix} 0 & \frac{1}{2} & 2 \\ \frac{1}{2} & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$, *d*) $\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$

13.27. *a*) $P = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}$, $P^TAP\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$; $r = 2$, $s = 0$.

b) $P = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}$, $P^TAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$; $r = 3$, $s = 1$.

c) $P = \begin{pmatrix} 1 & -1 & -1 & 26 \\ 0 & 1 & 3 & 13 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 7 \end{pmatrix}$, $P^TAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 469 \end{pmatrix}$; $r = 4$, $s = 2$.

CAPITULO 14

Operadores lineales en espacios con producto interno

14.1. INTRODUCCION

Este capítulo investiga el espacio $A(V)$ de operadores lineales T en un espacio con producto interno V. (Véase el Capítulo 6.) Así el cuerpo base K será bien el cuerpo real $\mathbf{R}$ o bien el cuerpo complejo $\mathbf{C}$. De hecho, se usará una terminología diferente para el caso real y para el caso complejo. Utilizaremos también el hecho de que el producto interno en el espacio euclídeo $\mathbf{R}^n$ puede definirse según

$$\langle u, v\rangle = u^T v$$

y el producto interno en el espacio euclídeo complejo $\mathbf{C}^n$, según

$$\langle u, v\rangle = u^T \bar{v}$$

donde u y v son vectores columna.

El lector debe revisar el material del Capítulo 6. En particular, debe familiarizarse con las nociones de norma (longitud), ortogonalidad y bases ortonormales.

Por último, debemos mencionar que el Capítulo 6 trataba principalmente de espacios reales con producto interno. Sin embargo, aquí supondremos que V es un espacio complejo con producto interno, a menos que se establezca o sobrentienda lo contrario.

14.2. OPERADORES ADJUNTOS

Comenzamos con la siguiente definición básica.

Definición: Se dice que un operador lineal T en un espacio con producto interno V tiene un *operador adjunto* T^* en V si $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ para todos los $u, v \in V$.

El próximo ejemplo muestra que el operador adjunto admite una descripción sencilla dentro del contexto de las aplicaciones matriciales.

EJEMPLO 14.1

a) Sea A una matriz n-cuadrada real vista como operador lineal en $\mathbf{R}^n$. Para todo los $u, v \in \mathbf{R}^n$,

$$\langle Au, v\rangle = (Au)^T v = u^T A^T v = \langle u, A^T v\rangle$$

De este modo, la matriz traspuesta A^T es el adjunto de A.

b) Sea B una matriz n-cuadrada compleja vista como operador lineal en $\mathbf{C}^n$. Para todos los $u, v \in \mathbf{C}^n$,

$$\langle Bu, v\rangle = (Bu)^T \bar{v} = u^T B^T \bar{v} = u^T \overline{\bar{B}^T} \bar{v} = u^T \bar{B}^* \bar{v} = \langle u, B^* v\rangle$$

La matriz traspuesta conjugada $B^* = \bar{B}^T$ es, pues, el adjunto de B.

Nota: La notación B^* se ha empleado para designar el adjunto de B y previamente para designar la traspuesta conjugada de B. El Ejemplo 14.1 muestra lo afortunado de tal coincidencia.

El teorema enunciado a continuación, demostrado en el Problema 14.4, es el resultado principal de la sección.

Teorema 14.1: Sea T un operador lineal en un espacio con producto interno de dimensión finita V sobre K. Entonces:

i) Existe un único operador lineal T^* en V tal que $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ para todos los $u, v \in V$. (Esto es, T tiene un adjunto T^*.)

ii) Si A es la representación matricial de T con respecto a cualquier base ortonormal $S = \{u_i\}$ de V, la representación matricial de T^* en la base S es la traspuesta conjugada A^* de A (o la traspuesta A^T si K es real).

Subrayamos que no existe una relación sencilla semejante entre las matrices que representan T y T^* si la base no es ortonormal. Vemos así una propiedad útil de las bases ortonormales. Hacemos énfasis también en que este teorema no es válido si V tiene dimensión infinita (Problema 14.31).

EJEMPLO 14.2. Sea T el operador lineal en $\mathbf{C}^3$ definido mediante

$$T(x, y, z) = (2x + iy,\ y - 5iz,\ x + (1 - i)y + 3z)$$

Hallemos una fórmula similar para el adjunto T^* de T. Nótese (Problema 10.1) que la matriz de T en la base usual de $\mathbf{C}^3$ es

$$[T] = \begin{pmatrix} 2 & i & 0 \\ 0 & 1 & -5i \\ 1 & 1-i & 3 \end{pmatrix}$$

Recordemos que la base usual es ortonormal. De esta manera, según el Teorema 14.1, la matriz de T^* en esta base es la traspuesta conjugada de $[T]$:

$$[T^*] = \begin{pmatrix} 2 & 0 & 1 \\ -i & 1 & 1+i \\ 0 & 5i & 3 \end{pmatrix}$$

En consecuencia,

$$T^*(x, y, z) = (2x + z, \ -ix + y + (1 + i)z, \ 5iy + 3z)$$

El siguiente teorema, demostrado en el Problema 14.5, resume algunas de las propiedades del adjunto.

Teorema 14.2: Sean T, T_1, T_2 operadores lineales en V y $k \in K$. Entonces:

i) $(T_1 + T_2)^* = T_1^* + T_2^*$

ii) $(kT)^* = \bar{k}T^*$

iii) $(T_1 T_2)^* = T_2^* T_1^*$

iv) $(T^*)^* = T$

Obsérvese la similitud entre el teorema precedente y el Teorema 3.3 relativo a las propiedades de la operación de trasposición sobre las matrices.

FUNCIONALES LINEALES Y ESPACIOS CON PRODUCTO INTERNO

Recordemos (Capítulo 12) que un funcional lineal ϕ en un espacio vectorial V es una aplicación lineal de V en el cuerpo base K. Esta subsección contiene un importante resultado (Teorema 14.3) que se emplea en la demostración del Teorema 14.1.

Sea V un espacio con producto interno. Cada $u \in V$ determina una aplicación $\hat{u}: V \to K$ definida por

$$\hat{u}(v) = \langle v, u \rangle$$

Ahora, para dos escalares cualesquiera a, $b \in K$ y dos vectores cualesquiera v_1, $v_2 \in V$,

$$\hat{u}(av_1 + bv_2) = \langle av_1 + bv_2, u \rangle = a\langle v_1, u \rangle + b\langle v_2, u \rangle = a\hat{u}(v_1) + b\hat{u}(v_2)$$

Es decir, $\hat{u}$ es un funcional lineal en V. El recíproco también es cierto para espacios de dimensión finita y es un importante teorema (demostrado en el Problema 14.3). A saber,

Teorema 14.3: Sea ϕ un funcional lineal en un espacio con producto interno de dimensión finita V. Existe un único vector $u \in V$ tal que $\phi(v) = \langle v, u \rangle$ para todo $v \in V$.

Señalamos que el teorema precedente no es válido para espacios de dimensión infinita (Problema 14.24), aunque se conocen algunos resultados generales que apuntan en esta dirección. (Uno de tales resultados es el famoso teorema de representación de Riesz.)

14.3. ANALOGIA ENTRE $A(V)$ Y C. OPERADORES ESPECIALES

Denotemos por $A(V)$ el álgebra de los operadores lineales en un espacio con producto interno de dimensión finita V. La aplicación adjunta $T \mapsto T^*$ en $A(V)$ presenta bastantes analogías con la aplicación de conjugación $z \mapsto \bar{z}$ en el cuerpo complejo **C**. Para ilustrar esta analogía identificamos en la Tabla 14.1 ciertas clases de operadores $T \in A(V)$ cuyo comportamiento bajo la aplicación adjunta imita el comportamiento bajo conjugación de clases familiares de números complejos.

Tabla 14-1

Clase de números complejos	*Comportamiento bajo conjugación*	*Clase de operadores en $A(V)$*	*Comportamiento bajo la aplicación adjunta*
Círculo unidad ($\|z\| = 1$)	$\bar{z} = 1/z$	Operadores ortogonales (caso real) Operadores unitarios (caso complejo)	$T^* = T^{-1}$
Eje real	$\bar{z} = z$	Operadores autoadjuntos También llamados: simétricos (caso real) hermíticos (caso complejo)	$T^* = T$
Eje imaginario	$\bar{z} = -z$	Operadores anti-autoadjuntos También llamados: antisimétricos (caso real) antihermíticos (caso complejo)	$T^* = -T$
Semieje real positivo $(0, \infty)$	$z = \bar{w}w,\ w \neq 0$	Operadores definidos positivos	$T = S^*S$ con S no singular

La analogía entre estas clases de operadores T y de números complejos z se refleja en el teorema que ahora enunciamos.

Teorema 14.4: Sea λ un valor propio de un operador lineal T en V.

i) Si $T^* = T^{-1}$ (es decir, si T es ortogonal o unitario), $|\lambda| = 1$.

ii) Si $T^* = T$ (es decir, si T es autoadjunto), λ es real.

iii) Si $T^* = -T$ (es decir, si T es anti-autoadjunto), λ es imaginario puro.

iv) Si $T = S^*S$ con S no singular (es decir, si T es definido positivo), λ es real y positivo.

Demostración. En cada caso, v será un vector propio no nulo de T perteneciente a λ, o sea, $T(v) = \lambda v$ con $v \neq 0$; por tanto, $\langle v, v\rangle$ será positivo.

Demostración de i): Probamos que $\lambda\bar{\lambda}\langle v, v\rangle = \langle v, v\rangle$:

$$\lambda\bar{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, I(v)\rangle = \langle v, v\rangle$$

Pero $\langle v, v\rangle \neq 0$, luego $\lambda\bar{\lambda} = 1$ y así $|\lambda| = 1$.

Demostración de ii): Probamos que $\lambda\langle v, v\rangle = \bar{\lambda}\langle v, v\rangle$:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle$$

Pero $\langle v, v\rangle \neq 0$, luego $\lambda = \bar{\lambda}$ y así λ es real.

Demostración de iii): Probamos que $\lambda\langle v, v\rangle = -\bar{\lambda}\langle v, v\rangle$:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, -T(v)\rangle = \langle v, -\lambda v\rangle = -\bar{\lambda}\langle v, v\rangle$$

Pero $\langle v, v\rangle \neq 0$, luego $\lambda = -\bar{\lambda}$ o $\bar{\lambda} = -\lambda$ y así λ es imaginario puro.

Demostración de iv): Nótese primero que $S(v) \neq 0$ porque S es no singular, luego $\langle S(v), S(v)\rangle$ es positivo. Problemos que $\lambda\langle v, v\rangle = \langle S(v), S(v)\rangle$:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle S^*S(v), v\rangle = \langle S(v), S(v)\rangle$$

Pero $\langle v, v\rangle$ y $\langle S(v), S(v)\rangle$ son positivos; por consiguiente, λ es positivo.

Nota: Cada uno de los operadores T anteriores conmuta con su adjunto, es decir, $TT^* = T^*T$. Tales operadores se denominan *operadores normales*.

14.4. OPERADORES AUTOADJUNTOS

Sea T un operador *autoadjunto* en un espacio con producto interno V, esto es, supongamos

$$T^* = T$$

(En caso de que T se defina mediante una matriz A, A será simétrica o hermítica, según sea real o compleja.) En virtud del Teorema 14.4, los valores propios de T son reales. Otra importante propiedad de T es la que sigue.

Teorema 14.5: Sea T un operador autoadjunto en V. Supongamos que u y v son vectores propios de T pertenecientes a valores propios distintos. Entonces u y v son ortogonales, o sea, $\langle u, v\rangle = 0$.

Demostración. Supongamos $T(u) = \lambda_1 u$ y $T(v) = \lambda_2 v$, donde $\lambda_1 \neq \lambda_2$. Probamos que $\lambda_1\langle u, v\rangle = \lambda_2\langle u, v\rangle$:

$$\lambda_1\langle u, v\rangle = \langle \lambda_1 u, v\rangle = \langle T(u), v\rangle = \langle u, T^*(v)\rangle = \langle u, T(v)\rangle =$$
$$= \langle u, \lambda_2 v\rangle = \bar{\lambda}_2\langle u, v\rangle = \lambda_2\langle u, v\rangle$$

(La cuarta igualdad utiliza que $T^* = T$ y la última que el valor propio λ_2 es real.) Dado que $\lambda_1 \neq \lambda_2$, conseguimos $\langle u, v\rangle = 0$. Queda, pues, demostrado el teorema.

14.5. OPERADORES ORTOGONALES Y UNITARIOS

Sea U un operador lineal en un espacio con producto interno de dimensión finita V. Recordemos que si

$$U^* = U^{-1} \quad \text{o equivalentemente} \quad UU^* = U^*U = I$$

se dice que U es ortogonal o unitario, según sea real o complejo el cuerpo subyacente. El próximo teorema, demostrado en el Problema 14.10, proporciona caracterizaciones alternativas de estos operadores.

Teorema 14.6: Las siguientes condiciones sobre un operador U son equivalentes:

i) $U^* = U^{-1}$, esto es, $UU^* = U^*U = I$.

ii) U preserva los productos internos, es decir, para todos los $v, w \in V$,

$$\langle U(v), U(w) \rangle = \langle v, w \rangle$$

iii) U preserva las longitudes, es decir, para todo $v \in V$, $\|U(v)\| = \|v\|$.

EJEMPLO 14.3

a) Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal que gira cada vector alrededor del eje z un ángulo θ fijo como se muestra en la Figura 14-1, esto es, el definido por

$$T(x, y, z) = (x \cos \theta - y \operatorname{sen} \theta, x \operatorname{sen} \theta + y \cos \theta, z)$$

Nótese que las longitudes (distancias desde el origen) se preservan bajo T. Siendo así, T es un operador ortogonal.

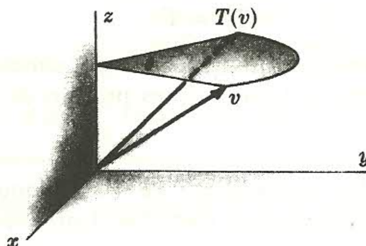


Figura 14-1.

b) Sea V el espacio l_2 del Ejemplo 6.3 *b*). Sea $T: V \to V$ el operador lineal definido por

$$T(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$$

Claramente, T preserva los productos internos y las longitudes. No obstante, T no es suprayectivo ya que, por ejemplo, $(1, 0, 0, \ldots)$ no pertenece a la imagen de T; por tanto, T no es invertible. Vemos, pues, que el Teorema 14.6 no es válido para espacios de dimensión infinita.

Un isomorfismo de un espacio con producto interno en otro es una aplicación biyectiva que preserva las tres operaciones básicas de un espacio con producto interno: suma de vectores, producto por un escalar y producto interno. De este modo, las aplicaciones precedentes (ortogonales y unitarias) pueden caracterizarse también como los isomorfismos de V en sí mismo. Nótese que una aplicación lineal tal, U, preserva asimismo las distancias, puesto que

$$\|U(v) - U(w)\| = \|U(v - w)\| = \|v - w\|$$

Por esa razón U se llama una *isometría*.

14.6. MATRICES ORTOGONALES Y UNITARIAS

Sea U un operador lineal en un espacio con producto interno V. De acuerdo con el Teorema 14.1, obtenemos el resultado que ahora enunciamos cuando el cuerpo base K es complejo.

Teorema 14.7A: Una matriz A con entradas complejas representa un operador unitario U (respecto a una base ortonormal) si y sólo si $A^* = A^{-1}$.

Por otra parte, si el cuerpo base K es real, $A^* = A^T$, luego tenemos el teorema correspondiente a espacios reales con producto interno siguiente.

Teorema 14.7B: Una matriz A con entradas reales representa un operador ortogonal U (respecto a una base ortonormal) si y sólo si $A^T = A^{-1}$.

Los teoremas anteriores motivan las definiciones escritas a continuación.

Definición: Una matriz compleja A para la que $A^* = A^{-1}$, o equivalentemente $AA^* = A^*A = I$, recibe el nombre de *matriz unitaria*.

Definición: Una matriz real A para la que $A^T = A^{-1}$, o equivalentemente $AA^T = A^TA = I$, recibe el nombre de *matriz ortogonal*.

Obsérvese que una matriz unitaria con entradas reales es ortogonal.

EJEMPLO 14.4. Supongamos que $A = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ es una matriz unitaria. En tal caso, $AA^* = I$ y por tanto

$$AA^* = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}\begin{pmatrix} \bar{a}_1 & \bar{b}_1 \\ \bar{a}_2 & \bar{b}_2 \end{pmatrix} = \begin{pmatrix} |a_1|^2 + |a_2|^2 & a_1\bar{b}_1 + a_2\bar{b}_2 \\ \bar{a}_1 b_1 + \bar{a}_2 b_2 & |b_1|^2 + |b_2|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

De esta manera,

$$|a_1|^2 + |a_2|^2 = 1 \qquad |b_1|^2 + |b_2|^2 = 1 \qquad \text{y} \qquad a_1\bar{b}_1 + a_2\bar{b}_2 = 0$$

En consecuencia, las filas de A forman un conjunto ortonormal. De forma similar, $A^*A = I$ obliga a las columnas de A a formar un conjunto ortonormal.

El resultado del ejemplo precedente es cierto en general; a saber,

Teorema 14.8: Las siguientes condiciones sobre una matriz A son equivalentes:

i) A es unitaria (ortogonal).
ii) Las filas de A forman un conjunto ortonormal.
iii) Las columnas de A forman un conjunto ortonormal.

14.7. CAMBIO DE BASE ORTONORMAL

En vista del papel especial de las bases ortonormales en la teoría de los espacios con producto interno, estamos interesados, naturalmente, en las propiedades de la matriz de cambio de base desde una de estas bases hasta otra. Disponemos del siguiente teorema, demostrado en el Problema 14.12.

Teorema 14.9: Sea $\{u_1, \ldots, u_n\}$ una base ortonormal de un espacio con producto interno V. La matriz de cambio de base desde $\{u_i\}$ hasta otra base ortonormal es unitaria (ortogonal). Recíprocamente, si $P = (a_{ij})$ es una matriz unitaria (ortogonal), la que sigue es una base ortonormal:

$$\{u_i' = a_{1i}u_1 + a_{2i}u_2 + \cdots + a_{ni}u_n : i = 1, \ldots, n\}$$

Recordemos que dos matrices A y B que representan el mismo operador lineal T son similares, o sea, $B = P^{-1}AP$, donde P es la matriz de cambio de base (no singular). Por otra parte, si V es un espacio con producto interno, suele interesarnos el caso en que P es unitaria (u ortogonal) como sugiere el Teorema 14.9. (Recuérdese que P es unitaria si $P^* = P^{-1}$, y ortogonal si $P^T = P^{-1}$.) Esto nos conduce a la definición que enseguida escribimos.

Definición: Dos matrices complejas A y B son *unitariamente equivalentes* si existe una matriz unitaria P para la cual $B = P^*AP$. Análogamente, dos matrices reales A y B son *ortogonalmente equivalentes* si existe una matriz ortogonal P para la cual $B = P^TAP$.

Nótese que las matrices ortogonalmente equivalentes son necesariamente congruentes.

14.8. OPERADORES POSITIVOS

Sea P un operador lineal en un espacio con producto interno V. Se dice que P es *positivo* (o *semidefinido*) si

$$P = S^*S \quad \text{para algún operador } S$$

y se dice que es *definido positivo* si, además, S es no singular. Los próximos teoremas dan caracterizaciones alternativas de estos operadores. (El Teorema 14.10A se demuestra en el Problema 14.21.)

Teorema 14.10A: Las siguientes condiciones sobre un operador P son equivalentes:

i) $P = T^2$ para algún operador autoadjunto T.
ii) P es positivo.
iii) P es autoadjunto y $\langle P(u), u\rangle \geq 0$ para todo $u \in V$.

El teorema correspondiente para operadores definidos positivos es

Teorema 14.10B: Las siguientes condiciones sobre un operador P son equivalentes:

i) $P = T^2$ para algún operador autoadjunto no singular T.
ii) P es definido positivo.
iii) P es autoadjunto y $\langle P(u), u\rangle > 0$ para todo $u \neq 0$ en V.

14.9. DIAGONALIZACION Y FORMAS CANONICAS EN ESPACIOS EUCLIDEOS

Sea T un operador lineal en un espacio con producto interno de dimensión finita V sobre K. La representación de T mediante una matriz diagonal depende de los vectores propios y valores

propios de T y por tanto de las raíces del polinomio característico $\Delta(t)$ de T. Ahora bien, $\Delta(t)$ siempre se factoriza en polinomios lineales sobre el cuerpo complejo $\mathbf{C}$, pero puede no tener polinomios lineales sobre el cuerpo real $\mathbf{R}$ como factores. La situación para espacios euclídeos (donde $K = \mathbf{R}$) es inherentemente distinta de la situación para espacios unitarios (donde $K = \mathbf{C}$); de ahí que los tratemos por separado. Investigamos los espacios euclídeos abajo (véase la demostración del Teorema 14.11 en el Problema 14.14) y los espacios unitarios en la siguiente sección.

Teorema 14.11: Sea T un operador simétrico (autoadjunto) en un espacio real con producto interno de dimensión finita V. Existe una base ortonormal de V constituida por vectores propios de T; esto es, T puede representarse por una matriz diagonal respecto a una base ortonormal.

Demos el enunciado correspondiente a matrices.

Teorema 14.11 (forma alternativa): Sea A una matriz real simétrica. Existe una matriz ortogonal P tal que $B = P^{-1}AP = P^{T}AP$ es diagonal.

Podemos elegir las columnas de la matriz ortogonal anterior P como vectores propios ortogonales normalizados de A; entonces las entradas diagonales de B son los correspondientes valores propios.

FORMA CANONICA PARA OPERADORES ORTOGONALES

Un operador ortogonal T no es necesariamente simétrico, por lo que puede no admitir una matriz simétrica como representación relativa a una base ortonormal. Sin embargo, tal operador T tiene una representación canónica simple, como se describe en el siguiente teorema, demostrado en el Problema 14.16.

Teorema 14.12: Sea T un operador ortogonal en un espacio real con producto interno V. Existe una base ortonormal con respecto a la cual T tiene la forma:

$$\begin{pmatrix}
1 & & & & & & & & & & & & \\
& 1 & & & & & & & & & & & \\
& & \ddots & & & & & & & & & & \\
& & & 1 & & & & & & & & & \\
& & & & -1 & & & & & & & & \\
& & & & & -1 & & & & & & & \\
& & & & & & \ddots & & & & & & \\
& & & & & & & -1 & & & & & \\
& & & & & & & & \cos\theta_1 & -\operatorname{sen}\theta_1 & & & \\
& & & & & & & & \operatorname{sen}\theta_1 & \cos\theta_1 & & & \\
& & & & & & & & & & \ddots & & \\
& & & & & & & & & & & \cos\theta_r & -\operatorname{sen}\theta_r \\
& & & & & & & & & & & \operatorname{sen}\theta_r & \cos\theta_r
\end{pmatrix}$$

El lector podrá reconocer cada uno de los bloques diagonales 2×2 precedentes como una rotación en el subespacio bidimensional correspondiente, y cada entrada diagonal -1 como una reflexión en el subespacio unidimensional correspondiente.

14.10. DIAGONALIZACION Y FORMAS CANONICAS EN ESPACIOS UNITARIOS

Presentamos ahora el teorema de diagonalización fundamental para espacios complejos con producto interno, es decir, para espacios unitarios. Recordemos que un operador T se dice normal si conmuta con su adjunto, o sea, si $TT^* = T^*T$. Análogamente, una matriz compleja A se dice normal si conmuta con su traspuesta conjugada, o sea, si $AA^* = A^*A$.

EJEMPLO 14.5. Sea $A = \begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix}$. Entonces

$$AA^* = \begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & 3-2i \end{pmatrix} = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 14 \end{pmatrix}$$
$$A^*A = \begin{pmatrix} 1 & -i \\ 1 & 3-2i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix} = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 14 \end{pmatrix}$$

De este modo, A es una matriz normal.

Es aplicable el teorema enunciado a continuación.

Teorema 14.13: Sea T un operador normal en un espacio complejo con producto interno de dimensión finita V. Existe una base ortonormal de V constituida por vectores propios de T; esto es, T puede representarse por una matriz diagonal respecto a una base ortonormal.

Demos el enunciado correspondiente para matrices.

Teorema 14.13 (forma alternativa): Sea A una matriz normal. Existe una matriz unitaria P tal que $B = P^{-1}AP = P^*AP$ es diagonal.

El siguiente teorema muestra que incluso los operadores no normales en espacios unitarios tienen una forma relativamente sencilla.

Teorema 14.14: Sea T un operador arbitrario en un espacio complejo con producto interno de dimensión finita V. Entonces T puede representarse por una matriz triangular respecto a una base ortonormal de V.

Teorema 14.14 (forma alternativa): Sea A una matriz compleja arbitraria. Existe una matriz unitaria P tal que $B = P^{-1}AP = P^*AP$ es triangular.

14.11. TEOREMA ESPECTRAL

El teorema espectral es una reformulación de los Teoremas 14.11 y 14.13 de diagonalización.

Teorema 14.15 (teorema espectral): Sea T un operador normal (simétrico) en un espacio complejo (real) con producto interno de dimensión finita V. Existen operadores lineales $E_1, \ldots, E_r$ en V y escalares $\lambda_1, \ldots, \lambda_r$ tales que

i) $T = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_r E_r$ iii) $E_1^2 = E_1, \ldots, E_r^2 = E_r$

ii) $E_1 + E_2 + \cdots + E_r = I$ iv) $E_i E_j = 0$ para $i \neq j$

Los operadores lineales anteriores $E_1, \ldots, E_r$ son *proyecciones* (o *proyectores*), en el sentido de que $E_i^2 = E_i$. Más aún, son *proyecciones ortogonales* porque tienen la propiedad adicional de que $E_i E_j = 0$ para $i \neq j$.

El próximo ejemplo muestra la relación entre una representación matricial diagonal y las proyecciones ortogonales correspondientes.

EJEMPLO 14.6. Consideremos una matriz diagonal, digamos $A = \begin{pmatrix} 2 & & & \\ & 3 & & \\ & & 3 & \\ & & & 5 \end{pmatrix}$. Sean

$$E_1 = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}$$

El lector puede comprobar que:

i) $A = 2E_1 + 3E_2 + 5E_3$, ii) $E_1 + E_2 + E_3 = I$, iii) $E_i^2 = E_i$, iv) $E_i E_j = 0$ para $i \neq j$.

PROBLEMAS RESUELTOS

ADJUNTOS

14.1. Hallar el adjunto de la matriz $B = \begin{pmatrix} 3-7i & 18 & 4+i \\ -7i & 6-i & 2-3i \\ 8+i & 7+9i & 6+3i \end{pmatrix}$.

La traspuesta conjugada es $B^* = \begin{pmatrix} 3+7i & 7i & 8-i \\ 18 & 6+i & 7-9i \\ 4-i & 2+3i & 6-3i \end{pmatrix}$.

14.2. Hallar el adjunto de cada operador lineal:

a) $F: \mathbf{R}^3 \to \mathbf{R}^3$ definido por $F(x, y, z) = (3x + 4y - 5z, 2x - 6y + 7z, 5x - 9y + z)$.

b) $G: \mathbf{C}^3 \to \mathbf{C}^3$ definido por

$$G(x, y, z) = (2x + (1 - i)y, (3 + 2i)x - 4iz, 2ix + (4 - 3i)y - 3z)$$

a) Primero hallamos la matriz A que representa F en la base usual de $\mathbf{R}^3$. (Recuérdese que las filas de A son los coeficientes de x, y, z.)

$$A = \begin{pmatrix} 3 & 4 & -5 \\ 2 & -6 & 7 \\ 5 & -9 & 1 \end{pmatrix}$$

Siendo $\mathbf{R}$ el cuerpo base, el adjunto F^* está representado por la traspuesta A^T de A. Construimos, pues,

$$A^T = \begin{pmatrix} 3 & 2 & 5 \\ 4 & -6 & -9 \\ -5 & 7 & 1 \end{pmatrix}$$

Entonces $F^*(x, y, z) = (3x + 2y + 5z, 4x - 6y - 9z, -5x + 7y + z)$.

b) Primero hallamos la matriz B que representa a G en la base usual de $\mathbf{C}^3$:

$$B = \begin{pmatrix} 2 & 1 - i & 0 \\ 3 + 2i & 0 & -4i \\ 2i & 4 - 3i & -3 \end{pmatrix}$$

Construimos la traspuesta conjugada B^* de B:

$$B^* = \begin{pmatrix} 2 & 3 - 2i & -2i \\ 1 + i & 0 & 4 + 3i \\ 0 & 4i & -3 \end{pmatrix}$$

De este modo, $G^*(x, y, z) = (2x + (3 - 2i)y - 2iz, (1 + i)x + (4 + 3i)z, 4iy - 3z)$.

14.3. Demostrar el Teorema 14.3.

Sea $\{w_1, \ldots, w_n\}$ una base ortonormal de V. Tomemos

$$u = \overline{\phi(w_1)}w_1 + \overline{\phi(w_2)}w_2 + \cdots + \overline{\phi(w_n)}w_n$$

Sea $\hat{u}$ el funcional lineal en V definido por $\hat{u}(v) = \langle v, u \rangle$ para todo $v \in V$. Para $i = 1, \ldots, n$,

$$\hat{u}(w_i) = \langle w_i, u \rangle = \langle w_i, \overline{\phi(w_1)}w_1 + \cdots + \overline{\phi(w_n)}w_n \rangle = \phi(w_i)$$

Como $\hat{u}$ y ϕ coinciden sobre cada vector de la base, $\hat{u} = \phi$.

Supongamos ahora que u' es otro vector en V para el cual $\phi(v) = \langle v, u' \rangle$ para todo $v \in V$. En ese caso, $\langle v, u \rangle = \langle v, u' \rangle$, o sea, $\langle v, u - u' \rangle = 0$. En particular, esto es cierto para $v = u - u'$ y así $\langle u - u', u - u' \rangle = 0$. Esto nos conduce a $u - u' = 0$ y $u = u'$. Tal vector u es, pues, único, como se pretendía probar.

14.4. Demostrar el Teorema 14.1.

Demostración de i):

Comenzamos definiendo la aplicación T^*. Sea v un elemento arbitrario pero fijo en V. La aplicación $u \mapsto \langle T(u), v \rangle$ es un funcional lineal en V. Por tanto, según el Teorema 14.3, existe un

único elemento $v' \in V$ tal que $\langle T(u), v\rangle = \langle u, v'\rangle$ para todo $u \in V$. Definimos $T^*: V \to V$ por $T^*(v) = v'$. Entonces $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ para todo par de vectores $u, v \in V$.

Probamos a continuación que T^* es lineal. Para cualesquiera $u, v_i \in V$ y $a, b \in K$,

$$\begin{aligned}\langle u, T^*(av_1 + bv_2)\rangle &= \langle T(u), av_1 + bv_2\rangle = \bar{a}\langle T(u), v_1\rangle + b\langle T(u), v_2\rangle = \\ &= \bar{a}\langle u, T^*(v_1)\rangle + \bar{b}\langle u, T^*(v_2)\rangle = \langle u, aT^*(v_1) + bT^*(v_2)\rangle\end{aligned}$$

Pero esto se cumple para todo $u \in V$, luego $T^*(av_1 + bv_2) = aT^*(v_1) + bT^*(v_2)$. De este modo, T^* es lineal.

Demostración de ii):

Las matrices $A = (a_{ij})$ y $B = (b_{ij})$ que representan T y T^*, respectivamente, en la base $S = \{u_i\}$ vienen dadas por $a_{ij} = \langle T(u_j), u_i\rangle$ y $b_{ij} = \langle T^*(u_j), u_i\rangle$ (véase el Problema 6.23). Por consiguiente,

$$b_{ij} = \langle T^*(u_j), u_i\rangle = \overline{\langle u_i, T^*(u_j)\rangle} = \overline{\langle T(u_i), u_j\rangle} = \overline{a_{ji}}$$

Así pues, $B = A^*$, como queríamos demostrar.

14.5. Demostrar el Teorema 14.2.

i) Para vectores cualesquiera $u, v \in V$,

$$\begin{aligned}\langle (T_1 + T_2)(u), v\rangle &= \langle T_1(u) + T_2(u), v\rangle = \langle T_1(u), v\rangle + \langle T_2(u), v\rangle = \\ &= \langle u, T_1^*(v)\rangle + \langle u, T_2^*(v)\rangle = \langle u, T_1^*(v) + T_2^*(v)\rangle = \\ &= \langle u, (T_1^* + T_2^*)(v)\rangle\end{aligned}$$

La unicidad del adjunto implica $(T_1 + T_2)^* = T_1^* + T_2^*$.

ii) Para vectores cualesquiera $u, v \in V$,

$$\langle (kT)(u), v\rangle = \langle kT(u), v\rangle = k\langle T(u), v\rangle = k\langle u, T^*(v)\rangle = \langle u, \bar{k}T^*(v)\rangle = \langle u, (\bar{k}T^*)(v)\rangle$$

La unicidad del adjunto implica $(kT)^* = \bar{k}T^*$.

iii) Para vectores cualesquiera $u, v \in V$,

$$\begin{aligned}\langle (T_1 T_2)(u), v\rangle &= \langle T_1(T_2(u)), v\rangle = \langle T_2(u), T_1^*(v)\rangle = \\ &= \langle u, T_2^*(T_1^*(v))\rangle = \langle u, (T_2^* T_1^*)(v)\rangle\end{aligned}$$

La unicidad del adjunto implica $(T_1 T_2)^* = T_2^* T_1^*$.

iv) Para vectores cualesquiera $u, v \in V$,

$$\langle T^*(u), v\rangle = \overline{\langle v, T^*(u)\rangle} = \overline{\langle T(v), u\rangle} = \langle u, T(v)\rangle$$

La unicidad del adjunto implica $(T^*)^* = T$.

14.6. Probar que: *a*) $I^* = I$, y *b*) $\mathbf{0}^* = \mathbf{0}$.

a) Para todos los $u, v \in V$, $\langle I(u), v\rangle = \langle u, v\rangle = \langle u, I(v)\rangle$, luego $I^* = I$.

b) Para todos los $u, v \in V$, $\langle \mathbf{0}(u), v\rangle = \langle 0, v\rangle = 0 = \langle u, 0\rangle = \langle u, \mathbf{0}(v)\rangle$, luego $\mathbf{0}^* = \mathbf{0}$.

14.7. Supóngase que T es invertible. Mostrar que $(T^{-1})^* = (T^*)^{-1}$.

$I = I^* = (TT^{-1})^* = (T^{-1})^* T^*$, luego $(T^{-1})^* = T^{*-1}$.

14.8. Sean T un operador lineal en V y W un subespacio T-invariante de V. Probar que $W^{\perp}$ es invariante bajo T^*.

Sea $u \in W^{\perp}$. Si $w \in W$, $T(w) \in W$ y así $\langle w, T^*(u)\rangle = \langle T(w), u\rangle = 0$. De esta manera, $T^*(u) \in W^{\perp}$ ya que es ortogonal a todo $w \in W$. Por tanto, $W^{\perp}$ es invariante bajo T^*.

14.9. Sea T un operador lineal en V. Mostrar que cada una de las siguientes condiciones implica $T = \mathbf{0}$:

i) $\langle T(u), v\rangle = 0$ para todos los $u, v \in V$.

ii) V es un espacio complejo y $\langle T(u), u\rangle = 0$ para todo $u \in V$.

iii) T es autoadjunto y $\langle T(u), u\rangle = 0$ para todo $u \in V$.

Dar un ejemplo de un operador T en un espacio real V para el que $\langle T(u), u\rangle = 0$ para todo $u \in V$ pero $T \neq \mathbf{0}$. [Así pues, ii) no se cumple necesariamente para un espacio real V.]

i) Tomemos $v = T(u)$. En tal caso, $\langle T(u), T(u)\rangle = 0$, luego $T(u) = 0$ para todo $u \in V$. En consecuencia, $T = \mathbf{0}$.

ii) Por hipótesis, $\langle T(v + w), v + w\rangle = 0$ para cualesquiera $v, w \in V$. Desarrollando y tomando $\langle T(v), v\rangle = 0$ y $\langle T(w), w\rangle = 0$,

$$\langle T(v), w\rangle + \langle T(w), v\rangle = 0 \qquad [1]$$

Nótese que w es arbitrario en [1]. Sustituyendo w por iw y usando $\langle T(v), iw\rangle = \bar{i}\langle T(v), w\rangle = -i\langle T(v), w\rangle$ y $\langle T(iw), v\rangle = \langle iT(w), v\rangle = i\langle T(w), v\rangle$,

$$-i\langle T(v), w\rangle + i\langle T(w), v\rangle = 0$$

Dividiendo por i y sumando el resulado a [1] obtenemos $\langle T(w), v\rangle = 0$ para cualesquiera $v, w \in V$. Por i), $T = \mathbf{0}$.

iii) Por ii) el resultado se cumple en el caso complejo; de ahí que sólo necesitemos considerar el caso real. Desarrollando $\langle T(v + w), v + w\rangle = 0$ llegamos nuevamente a [1]. Puesto que T es autoadjunto y el espacio es real, tenemos $\langle T(w), v\rangle = \langle w, T(v)\rangle = \langle T(v), w\rangle$. Sustituyéndolo en [1] obtenemos $\langle T(v), w\rangle = 0$ para cualesquiera $v, w \in V$. De acuerdo con i), $T = \mathbf{0}$.

Para nuestro ejemplo, consideremos el operador lineal T en $\mathbf{R}^2$ definido por $T(x, y) = (y, -x)$. Entonces $\langle T(u), u\rangle = 0$ para todo $u \in V$, pero $T \neq \mathbf{0}$.

OPERADORES Y MATRICES ORTOGONALES Y UNITARIOS

14.10. Demostrar el Teorema 14.6.

Supongamos que se cumple i). Para todo par de vectores $v, w \in V$,

$$\langle U(v), U(w)\rangle = \langle v, U^*U(w)\rangle = \langle v, I(w)\rangle = \langle v, w\rangle$$

De este modo, i) implica ii). Ahora bien, si se verifica ii),

$$\|U(v)\| = \sqrt{\langle U(v), U(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$$

Por consiguiente, ii) implica iii). Queda probar que iii) implica i).

Supongamos que se cumple iii). En ese caso, para todo $v \in V$,

$$\langle U^*U(v), v\rangle = \langle U(v), U(v)\rangle = \langle v, v\rangle = \langle I(v), v\rangle$$

Por tanto, $\langle (U^*U - I)(v), v\rangle = 0$ para todo $v \in V$. Pero $U^*U - I$ es autoadjunto (demuéstrese); entonces, por el Problema 14.9, tenemos $U^*U - I = 0$, por lo que $U^*U = I$. Así que $U^* = U^{-1}$, como se había anunciado.

14.11. Sean U un operador unitario (ortogonal) en V y W un subespacio invariante bajo U. Mostrar que $W^\perp$ también es invariante bajo U.

Como U es no singular, $U(W) = W$; esto es, para cualquier $w \in W$ existe $w' \in W$ tal que $U(w') = w$. Sea ahora $v \in W^\perp$. Para cualquier $w \in W$,

$$\langle U(v), w\rangle = \langle U(v), U(w')\rangle = \langle v, w'\rangle = 0$$

De esta manera, $U(v)$ pertenece a $W^\perp$. Por tanto, $W^\perp$ es invariante bajo U.

14.12. Demostrar el Teorema 14.9.

Supongamos que $\{v_i\}$ es otra base ortonormal y que

$$v_i = b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n \qquad i = 1, \ldots, n \qquad [1]$$

Dado que $\{v_i\}$ es ortonormal,

$$\delta_{ij} = \langle v_i, v_j\rangle = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}} \qquad [2]$$

Sea $B = (b_{ij})$ la matriz de coeficientes en [1]. (Entonces B^T es la matriz de cambio de base desde $\{u_i\}$ hasta $\{v_i\}$.) En tal caso, $BB^* = (c_{ij})$, donde $c_{ij} = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}}$. Por [2], $c_{ij} = \delta_{ij}$ y por consiguiente $BB^* = I$. De acuerdo con esto, B y por tanto B^T son unitarias.

Resta probar que $\{u_i'\}$ es ortonormal. Según el Problema 6.23,

$$\langle u_i', u_j'\rangle = a_{1i}\overline{a_{1j}} + a_{2i}\overline{a_{2j}} + \cdots + a_{ni}\overline{a_{nj}} = \langle C_i, C_j\rangle$$

donde C_i denota la columna i-ésima de la matriz unitaria (ortogonal) $P = (a_{ij})$. Siendo P unitaria (ortogonal), sus columnas forman un conjunto ortonormal, luego $\langle u_i', u_j'\rangle = \langle C_i, C_j\rangle = \delta_{ij}$. Así pues, $\{u_i'\}$ es una base ortonormal.

OPERADORES SIMETRICOS Y FORMAS CANONICAS EN ESPACIOS EUCLIDEOS

14.13. Sea T un operador simétrico. Probar que: *a*) el polinomio característico $\Delta(t)$ de T es producto de polinomios lineales (sobre **R**); *b*) T tiene un vector propio no nulo.

a) Sea A la matriz que representa a T respecto a una base ortonormal de V; entonces $A = A^T$. Sea $\Delta(t)$ el polinomio característico de A. Viendo A como un operador autoadjunto complejo, A sólo tiene valores propios reales en virtud del Teorema 14.4. Siendo así,

$$\Delta(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n)$$

donde los λ_i son todos reales. En otras palabras, $\Delta(t)$ es producto de polinomios lineales sobre **R**.

b) Por *a*), T tiene al menos un valor propio (real). De aquí que T tenga un vector propio no nulo.

14.14. Demostrar el Teorema 14.11.

Se demuestra por inducción en la dimensión de V. Si dim $V = 1$, el teorema se cumple trivialmente. Supongamos ahora dim $V = n > 1$. Según el Problema 14.13, existe un vector propio no nulo v_1 de T. Sean W el espacio generado por v_1 y u_1 un vector unitario en W, por ejemplo, $u_1 = v_1/\|v_1\|$.

Como v_1 es un vector propio de T, el subespacio W de V es invariante bajo T. Por el Problema 14.8, $W^\perp$ es invariante bajo $T^* = T$. De este modo, la restricción $\hat{T}$ de T a $W^\perp$ es un operador simétrico. De acuerdo con el Teorema 6.4, $V = W \oplus W^\perp$. Por tanto, dim $W^\perp = n - 1$ porque dim $W = 1$. Por inducción, existe una base ortonormal $\{u_2, \ldots, u_n\}$ de $W^\perp$ que consiste en vectores propios de $\hat{T}$ y por ende de T. Pero $\langle u_1, u_i \rangle = 0$ para $i = 2, \ldots, n$ puesto que $u_i \in W^\perp$. En consecuencia, $\{u_1, u_2, \ldots, u_n\}$ es un conjunto ortonormal constituido por vectores propios de T. Queda, pues, demostrado el teorema.

14.15. Encontrar un cambio de coordenadas ortogonal que diagonalice la forma cuadrática real definida como $q(x, y) = 2x^2 + 2xy + 2y^2$.

Empezamos hallando la matriz simétrica A que representa a q y después su polinomio característico $\Delta(t)$:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{y} \quad \Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 \\ -1 & t-2 \end{vmatrix} = (t-1)(t-3)$$

Los valores propios de A son 1 y 3, luego la forma diagonal de q es

$$q(x', y') = x'^2 + 3y'^2$$

Hallamos la transformación de coordenadas asociada obteniendo un conjunto ortonormal de vectores propios de A correspondientes.

Hacemos $t = 1$ en la matriz $tI - A$ para llegar al sistema homogéneo.

$$-x - y = 0 \qquad -x - y = 0$$

Una solución no nula es $v_1 = (1, -1)$. Ahora hacemos $t = 3$ en la matriz $tI - A$ para encontrar el sistema homogéneo

$$x - y = 0 \qquad -x + y = 0$$

Una solución no nula es $v_2 = (1, 1)$. Como cabía esperar del Teorema 14.5, v_1 y v_2 son ortogonales. Normalizamos v_1 y v_2 para obtener la base ortonormal.

$$\{u_1 = (1/\sqrt{2}, -1/\sqrt{2}), u_2 = (1/\sqrt{2}, 1/\sqrt{2})\}$$

La matriz de transición P y la transformación de coordenadas requerida son las siguientes:

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{o} \quad \begin{matrix} x = (x' + y')/\sqrt{2} \\ y = (-x' + y')/\sqrt{2} \end{matrix}$$

Nótese que las columnas de P son u_1 y u_2. También podemos expresar x' e y' en términos de x e y utilizando $P^{-1} = P^T$; esto es,

$$x' = (x - y)/\sqrt{2} \qquad y' = (x + y)/\sqrt{2}$$

14.16. Demostrar el Teorema 14.12.

Sea $S = T + T^{-1} = T + T^*$. Entonces, $S^* = (T + T^*)^* = T^* + T = S$. Así que S es un operador simétrico en V. Por el Teorema 14.11, existe una base ortonormal de V formada por vectores

propios de S. Si $\lambda_1, \ldots, \lambda_m$ denotan los valores propios distintos de S, V puede descomponerse en la suma directa $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$, donde V_i consiste en los vectores propios de S pertenecientes a λ_i. Afirmamos que cada V_i es invariante bajo T. En efecto, supongamos $v \in V_i$; en ese caso, $S(v) = \lambda_i v$ y

$$S(T(v)) = (T + T^{-1})T(v) = T(T + T^{-1})(v) = TS(v) = T(\lambda_i v) = \lambda_i T(v)$$

Esto es, $T(v) \in V_i$. Por consiguiente, V_i es invariante bajo T. Puesto que los V_i son ortogonales entre sí, podemos limitarnos a investigar cómo actúa T sobre cada V_i individual.

En un V_i dado, $(T + T^{-1})v = S(v) = \lambda_i v$. Multiplicando por T,

$$(T^2 - \lambda_i T + I)(v) = 0$$

Consideramos los casos $\lambda_i = \pm 2$ y $\lambda_i \neq \pm 2$ por separado. Si $\lambda_i = \pm 2$, $(T \pm I)^2(v) = 0$, lo que conduce a $(T \pm I)(v) = 0$ o $T(v) = \pm v$. De este modo, T restringido a este V_i es I o $-I$.

Si $\lambda_i \neq \pm 2$, T no tiene vectores propios en V_i ya que, según el Teorema 14.4, los únicos valores propios de T son 1 o -1. En consecuencia, para $v \neq 0$ los vectores v y $T(v)$ son linealmente independientes. Sea W el subespacio generado por v y $T(v)$. Entonces W es invariante bajo T porque

$$T(T(v)) = T^2(v) = \lambda_i T(v) - v$$

En virtud del Teorema 6.4, $V_i = W \oplus W^\perp$. Además, por el Problema 14.8, $W^\perp$ es también invariante bajo T. Siendo así, podemos descomponer V_i en la suma directa de subespacios bidimensionales W_j ortogonales entre sí e invariantes bajo T. Podemos, pues, limitarnos a investigar cómo actúa T sobre cada W_j individual.

Dado que $T^2 - \lambda_i T + I = 0$, el polinomio característico $\Delta(t)$ de T actuando sobre W_j es $\Delta(t) = t^2 - \lambda_i t + 1$. Por tanto, el determinante de T es 1, el término constante en $\Delta(t)$. De acuerdo con el Teorema 4.6, la matriz A que representa a T actuando sobre W_j respecto a una base ortonormal de W_j debe ser de la forma

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

La unión de las bases de los W_j proporciona una base ortonormal de V_i, y la de las bases de los V_i una base ortonormal de V en la cual la matriz que representa a T tiene la forma deseada.

OPERADORES NORMALES Y FORMAS CANONICAS EN ESPACIOS UNITARIOS

14.17. Determinar qué matriz es normal: *a)* $A = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$, *b)* $B = \begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix}$.

a) $AA^* = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix}$, $\quad A^*A = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$.

Como $AA^* \neq A^*A$, la matriz A no es normal.

b) $BB^* = \begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -i & 2-i \end{pmatrix} = \begin{pmatrix} 2 & 2+2i \\ 2-2i & 6 \end{pmatrix}$,

$B^*B = \begin{pmatrix} 1 & 1 \\ -i & 2-i \end{pmatrix}\begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix} = \begin{pmatrix} 2 & 2+2i \\ 2-2i & 6 \end{pmatrix}$.

Como $BB^* = B^*B$, la matriz B es normal.

14.18. Sea T un operador normal. Demostrar:

a) $T(v) = 0$ si y sólo si $T^*(v) = 0$.

b) $T - \lambda I$ es normal.

c) Si $T(v) = \lambda v$, necesariamente $T^*(v) = \bar{\lambda} v$, luego todo vector propio de T lo es también de T^*.

d) Si $T(v) = \lambda_1 v$ y $T(w) = \lambda_2 w$ con $\lambda_1 \neq \lambda_2$, necesariamente $\langle v, w\rangle = 0$; o sea, vectores propios de T correspondientes a valores propios distintos son ortogonales.

a) Probemos que $\langle T(v), T(v)\rangle = \langle T^*(v), T^*(v)\rangle$:

$$\langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, TT^*(v)\rangle = \langle T^*(v), T^*(v)\rangle$$

Por consiguiente, de acuerdo con $[I_3]$, $T(v) = 0$ si y sólo si $T^*(v) = 0$.

b) Probemos que $T - \lambda I$ conmuta con su adjunto:

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \bar{\lambda} I) = TT^* - \lambda T^* - \bar{\lambda} T + \lambda\bar{\lambda} I = \\ &= T^*T - \bar{\lambda} T - \lambda T^* + \bar{\lambda}\lambda I = (T^* - \bar{\lambda} I)(T - \lambda I) = \\ &= (T - \lambda I)^*(T - \lambda I)\end{aligned}$$

Así pues, $T - \lambda I$ es normal.

c) Si $T(v) = \lambda v$, $(T - \lambda I)(v) = 0$. Ahora bien, $T - \lambda I$ es normal según *b*); por tanto, por *a*), $(T - \lambda I)^*(v) = 0$. Esto es, $(T^* - \bar{\lambda} I)(v) = 0$, luego $T^*(v) = \bar{\lambda} v$.

d) Probamos que $\lambda_1\langle v, w\rangle = \lambda_2\langle v, w\rangle$:

$$\lambda_1\langle v, w\rangle = \langle \lambda_1 v, w\rangle = \langle T(v), w\rangle = \langle v, T^*(w)\rangle = \langle v, \bar{\lambda}_2 w\rangle = \lambda_2\langle v, w\rangle$$

Pero $\lambda_1 \neq \lambda_2$, luego $\langle v, w\rangle = 0$.

14.19. Demostrar el Teorema 14.13.

Se demuestra por inducción en la dimensión de V. Si $\dim V = 1$, el teorema se cumple trivialmente. Supongamos ahora $\dim V = n > 1$. Como V es un espacio vectorial complejo, T tiene al menos un valor propio y por tanto un vector propio no nulo v. Sean W el subespacio generado por v y u_1 un vector unitario en W.

Puesto que v es un vector propio de T, el subespacio W es invariante bajo T. No obstante, v es asimismo un vector propio de T^* por el Problema 14.18; de aquí que W sea también invariante bajo T^*. De acuerdo con el Problema 14.8, $W^\perp$ es invariante bajo $T^{**} = T$. El resto de la demostración es idéntico a la última parte de la del Teorema 14.11 (Problema 14.14).

14.20. Demostrar el Teorema 14.14.

Se demuestra por inducción en la dimensión de V. Si $\dim V = 1$, el teorema se cumple trivialmente. Supongamos ahora $\dim V = n > 1$. Como V es un espacio vectorial complejo, T tiene al menos un valor propio y por ende un vector propio no nulo v. Sean W el subespacio generado por v y u_1 un vector unitario en W. Entonces u_1 es un vector propio de T y, digamos, $T(u_i) = a_{11}u_1$.

Según el Teorema 6.4, $V = W \oplus W^\perp$. Denotemos por E la proyección ortogonal de V en $W^\perp$. Claramente $W^\perp$ es invariante bajo el operador ET. Por inducción, existe una base ortonormal $\{u_2, \ldots, u_n\}$ de $W^\perp$ tal que para $i = 2, \ldots, n$,

$$ET(u_i) = a_{i2}u_2 + a_{i3}u_3 + \cdots + a_{ii}u_i$$

(Nótese que $\{u_1, u_2, \ldots, u_n\}$ es una base ortonormal de V.) Pero E es la proyección ortogonal de V sobre $W^\perp$, luego debe cumplirse

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{ii}u_i$$

para $i = 2, \ldots, n$. Esto, con $T(u_1) = a_{11}u_1$, nos da el resultado deseado.

PROBLEMAS VARIOS

14.21. Demostrar el Teorema 14.10A.

Supongamos que se cumple i), esto es, que $P = T^2$, donde $T = T^*$. En ese caso, $P = TT = T^*T$ y así i) implica ii). Supongamos ahora que se verifica ii). Entonces $P^* = (S^*S)^* = S^*S^{**} = S^*S = P$ y así P es autoadjunto. Además,

$$\langle P(u), u\rangle = \langle S^*S(u), u\rangle = \langle S(u), S(u)\rangle \geq 0$$

De este modo, ii) implica iii), por lo que sólo queda probar que iii) implica i).

Supongamos que se cumple iii). Dado que P es autoadjunto, existe una base ortonormal $\{u_1, \ldots, u_n\}$ de V consistente en vectores propios de P; sea, por ejemplo, $P(u_i) = \lambda_i u_i$. De acuerdo con el Teorema 14.4, los λ_i son reales. Usando iii) mostramos que los λ_i son no negativos. Tenemos, para cada i,

$$0 \leq \langle P(u_i), u_i\rangle = \langle \lambda_i u_i, u_i\rangle = \lambda_i\langle u_i, u_i\rangle$$

Así pues, $\langle u_i, u_i\rangle \geq 0$ hace forzoso $\lambda_i \geq 0$, como se pretendía. En consecuencia, $\sqrt{\lambda_i}$ es un número real. Sea T el operador lineal definido según

$$T(u_i) = \sqrt{\lambda_i}u_i \qquad \text{para } i = 1, \ldots, n$$

Como T está representado por una matriz real diagonal respecto a la base ortonormal $\{u_i\}$, T es autoadjunto. Más aún, para cada i,

$$T^2(u_i) = T(\sqrt{\lambda_i}u_i) = \sqrt{\lambda_i}\,T(u_i) = \sqrt{\lambda_i}\sqrt{\lambda_i}u_i = \lambda_i u_i = P(u_i)$$

Puesto que T^2 y P coinciden sobre una base de V, $P = T^2$. Queda, pues, demostrado el teorema.

Nota: El operador T precedente es el único operador positivo tal que $P = T^2$; se llama la *raíz cuadrada positiva* de P.

14.22. Mostrar que cualquier operador T es suma de un operador autoadjunto y uno anti-autoadjunto.

Tomemos $S = \frac{1}{2}(T + T^*)$ y $U = \frac{1}{2}(T - T^*)$. Entonces $T = S + U$, donde

$$S^* = (\tfrac{1}{2}(T + T^*))^* = \tfrac{1}{2}(T^* + T^{**}) = \tfrac{1}{2}(T^* + T) = S$$

y

$$U^* = (\tfrac{1}{2}(T - T^*))^* = \tfrac{1}{2}(T^* - T) = -\tfrac{1}{2}(T - T^*) = -U$$

es decir, S es autoadjunto y U anti-autoadjunto.

14.23. Demostrar la siguiente afirmación: Sea T un operador lineal arbitrario en un espacio con producto interno de dimensión finita V. Entonces T es producto de un operador unitario (ortogonal) U y un único operador positivo P, esto es, $T = UP$. Además, si T es invertible, U también está unívocamente determinado.

Por el Teorema 14.10, T^*T es un operador positivo y por tanto existe un (único) operador positivo P tal que $P^2 = T^*T$ (Problema 14.43). Obsérvese que

$$\|P(v)\|^2 = \langle P(v), P(v)\rangle = \langle P^2(v), v\rangle = \langle T^*T(v), v\rangle = \langle T(v), T(v)\rangle = \|T(v)\|^2 \quad [1]$$

Consideramos ahora por separado las situaciones en que T es invertible y no invertible.

Si T es invertible, tomamos $\hat{U} = PT^{-1}$. Probemos que $\hat{U}$ es unitario:

$$\hat{U}^* = (PT^{-1})^* = T^{-1*}P^* = (T^*)^{-1}P \qquad \text{y} \qquad \hat{U}^*\hat{U} = (T^*)^{-1}PPT^{-1} = (T^*)^{-1}T^*TT^{-1} = I$$

De este modo, $\hat{U}$ es unitario. A continuación tomamos $U = \hat{U}^{-1}$. En tal caso, U es a su vez unitario y $T = UP$, como se pedía.

Para demostrar la unicidad suponemos $T = U_0P_0$ con U_0 unitario y P_0 positivo. Entonces

$$T^*T = P_0^*U_0^*U_0P_0 = P_0IP_0 = P_0^2$$

Pero la raíz cuadrada de T^*T es única (Problema 14.43), luego $P_0 = P$. (Nótese que la invertibilidad de T no se ha utilizado para demostrar la unicidad de P.) Ahora, si T es invertible, necesariamente lo es P por [1]. Multiplicar por la derecha $U_0P = UP$ por P^{-1} proporciona $U_0 = U$. Así U es también único cuando T es invertible.

Supongamos ahora que T no es invertible. Sea W la imagen de P, o sea, $W = \text{Im } P$. Definimos $U_1 : W \to V$ según

$$U_1(w) = T(v) \qquad \text{donde} \qquad P(v) = w \quad [2]$$

Debemos probar que U_1 está bien definido, esto es, que $P(v) = P(v')$ implica $T(v) = T(v')$. Esto deriva del hecho de que $P(v - v') = 0$ es equivalente a $\|P(v - v')\| = 0$ que hace forzoso $\|T(v - v')\| = 0$ por [1]. Siendo así, U_1 está bien definido. A continuación definimos $U_2 : W^\perp \to V$. Nótese que por [1] P y T tienen el mismo núcleo. Por consiguiente, las imágenes de P y T tienen la misma dimensión, es decir, dim (Im P) = dim W = dim (Im T). En consecuencia, $W^\perp$ e (Im $T)^\perp$ también tienen la misma dimensión. Sea U_2 cualquier isomorfismo entre $W^\perp$ e (Im $T)^\perp$.

Tomamos ahora $U = U_1 \oplus U_2$. [Aquí U se define como sigue: si $v \in V$ y $v = w + w'$, donde $w \in W$, $w' \in W^\perp$, $U(v) = U_1(w) + U_2(w')$.] Ahora bien, U es lineal (Problema 14.68) y, si $v \in V$ y $P(v) = w$, por [2],

$$T(v) = U_1(w) = U(w) = UP(v)$$

De esta manera, $T = UP$ como se pedía.

Falta probar que U es unitario. Todo vector $x \in V$ puede escribirse como $x = P(v) + w'$ con $w' \in W^\perp$. Entonces $U(x) = UP(v) + U_2(w') = T(v) + U_2(w')$, donde $\langle T(v), U_2(w')\rangle = 0$ por definición de U_2. Asimismo, $\langle T(v), T(v)\rangle = \langle P(v), P(v)\rangle$ por [1]. Tenemos, pues,

$$\begin{aligned}\langle U(x), U(x)\rangle &= \langle T(v) + U_2(w'), T(v) + U_2(w')\rangle = \\ &= \langle T(v), T(v)\rangle + \langle U_2(w'), U_2(w')\rangle = \\ &= \langle P(v), P(v)\rangle + \langle w', w'\rangle = \langle P(v) + w', P(v) + w'\rangle = \\ &= \langle x, x\rangle\end{aligned}$$

[Hemos usado también el hecho de que $\langle P(v), w'\rangle = 0$.] Por tanto, U es unitario y queda demostrado el teorema.

14.24. Sea V el espacio vectorial de los polinomios sobre **R** con producto interno definido mediante

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt.$$

Dar un ejemplo de funcional lineal ϕ en V para el que no se cumpla el Teorema 14.3, o sea, tal que no exista un polinomio $h(t)$ para el cual $\phi(f) = \langle f, h\rangle$ para todo $f \in V$.

Definamos $\phi: V \to \mathbf{R}$ por $\phi(f) = f(0)$, es decir, ϕ evalúa $f(t)$ en 0 y por tanto aplica $f(t)$ en su término constante. Supongamos que existe un polinomio $h(t)$ para el cual

$$\phi(f) = f(0) = \int_0^1 f(t)h(t)\,dt \qquad [1]$$

para todo polinomio $f(t)$. Observemos que ϕ aplica el polinomio $tf(t)$ en 0; de aquí, por [1],

$$\int_0^1 tf(t)h(t)\,dt = 0 \qquad [2]$$

para todo polinomio $f(t)$. En particular, [2] debe verificarse para $f(t) = th(t)$, esto es,

$$\int_0^1 t^2h^2(t)\,dt = 0$$

Esta integral hace que $h(t)$ sea el polinomio cero, luego $\phi(f) = \langle f, h\rangle = \langle f, 0\rangle = 0$ para todo polinomio $f(t)$. Esto contradice el hecho de que ϕ no es el funcional cero; por consiguiente, no existe el polinomio $h(t)$.

PROBLEMAS SUPLEMENTARIOS

OPERADORES ADJUNTOS

14.25. Hallar la adjunta de cada una de las siguientes matrices:

a) $A = \begin{pmatrix} 5-2i & 3+7i \\ 4-6i & 8+3i \end{pmatrix}$, *b)* $B = \begin{pmatrix} 3 & 5i \\ i & -2i \end{pmatrix}$, *c)* $C = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$

14.26. Defínase $T: \mathbf{R}^3 \to \mathbf{R}^3$ por $T(x, y, z) = (x + 2y, 3x - 4z, y)$. Hallar $T^*(x, y, z)$.

14.27. Defínase $T: \mathbf{C}^3 \to \mathbf{C}^3$ por

$$T(x, y, z) = (ix + (2+3i)y, 3x + (3-i)z, (2-5i)y + iz)$$

Hallar $T^*(x, y, z)$.

14.28. Para cada uno de los siguientes funcionales lineales ϕ en V, encontrar un vector $u \in V$ tal que $\phi(v) = \langle v, u\rangle$ para todo $v \in V$:

a) $\phi: \mathbf{R}^3 \to \mathbf{R}$ definido por $\phi(x, y, z) = x + 2y - 3z$.

b) $\phi: \mathbf{C}^3 \to \mathbf{C}$ definido por $\phi(x, y, z) = ix + (2+3i)y + (1-2i)z$.

c) $\phi: V \to \mathbf{R}$ definido por $\phi(f) = f(1)$, donde V es el espacio vectorial del Problema 14.24.

14.29. Supóngase que V tiene dimensión finita. Demostrar que la imagen de T^* es el complemento ortogonal del núcleo de T, o sea, Im $T^* = (\text{Ker } T)^\perp$. Por tanto, rango $T =$ rango T^*.

14.30. Mostrar que $T^*T = 0$ implica $T = 0$.

14.31. Sea V el espacio vectorial de los polinomios sobre $\mathbf{R}$ con producto interno definido según $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Sea $\mathbf{D}$ el operador derivada en V, es decir, $\mathbf{D}(f) = df/dt$. Probar que no hay ningún operador $\mathbf{D}^*$ en V tal que $\langle \mathbf{D}(f), g\rangle = \langle f, \mathbf{D}^*(g)\rangle$ para todos los $f, g \in V$. Esto es, $\mathbf{D}$ no tiene adjunto.

OPERADORES Y MATRICES UNITARIOS Y ORTOGONALES

14.32. Hallar una matriz unitaria (ortogonal) cuya primera fila sea:

a) $(2/\sqrt{13}, 3/\sqrt{13})$, *b*) un múltiplo de $(1, 1 - i)$, *c*) $(\frac{1}{2}, \frac{1}{2}i, \frac{1}{2} - \frac{1}{2}i)$.

14.33. Demostrar la siguiente afirmación: Los productos y las inversas de las matrices ortogonales son ortogonales. (Las matrices ortogonales forman, pues, un grupo bajo el producto, llamado el *grupo ortogonal*.)

14.34. Demostrar la siguiente afirmación: Los productos y las inversas de las matrices unitarias son unitarias. (Las matrices unitarias forman, pues, un grupo bajo el producto, llamado el *grupo unitario*.)

14.35. Mostrar que si una matriz ortogonal (unitaria) es triangular, es necesariamente diagonal.

14.36. Recuérdese que dos matrices complejas A y B son unitariamente equivalentes si existe una matriz unitaria P tal que $B = P^*AP$. Probar que ésta es una relación de equivalencia.

14.37. Recuérdese que dos matrices reales A y B son ortogonalmente equivalentes si existe una matriz ortogonal P tal que $B = P^TAP$. Probar que ésta es una relación de equivalencia.

14.38. Sea W un subespacio de V. Para cualquier $v \in V$, sea $v = w + w'$ con $w \in W$, $w' \in W^\perp$. (Tal suma es única porque $V = W \oplus W^\perp$.) Defínase $T: V \to V$ por $T(v) = w - w'$. Mostrar que T es un operador autoadjunto unitario en V.

14.39. Sea V un espacio con producto interno y supóngase que $U: V \to V$ (no necesariamente lineal) es suprayectivo y preserva los productos internos, o sea, $\langle U(v), U(w)\rangle = \langle v, w\rangle$ para todo par de vectores $v, w \in V$. Demostrar que U es lineal y por tanto unitario.

OPERADORES POSITIVOS Y DEFINIDOS POSITIVOS

14.40. Probar que la suma de dos operadores positivos (definidos positivos) es positiva (definida positiva).

14.41. Sea T un operador lineal en V y defínase $f: V \times V \to K$ por $f(u, v) = \langle T(u), v\rangle$. Probar que f es a su vez un producto interno en V si y sólo si T es definido positivo.

14.42. Supóngase que E es una proyección ortogonal sobre algún subespacio W de V. Demostrar que $kI + E$ es positivo (definido positivo) si $k \geq 0$ ($k > 0$).

14.43. Considérese el operador T definido por $T(u_i) = \sqrt{\lambda_i}u_i$, $i = 1, \ldots, n$, en la demostración del Teorema 14.10A. Mostrar que T es positivo y que es el único operador positivo para el cual $T^2 = P$.

14.44. Supóngase que P es positivo y unitario. Demostrar que $P = I$.

14.45. Determinar cuáles de las siguientes matrices son positivas (definidas positivas):

$$\text{i) } \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{ii) } \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \quad \text{iii) } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{iv) } \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{v) } \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{vi) } \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

14.46. Demostrar que una matriz compleja 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es positiva si y sólo si: i) $A = A^*$, ii) a, d y $ad - bc$ son números reales no negativos.

14.47. Demostrar que una matriz diagonal A es positiva (definida positiva) si y sólo si toda entrada diagonal es un número real no negativo (positivo).

OPERADORES AUTOADJUNTOS Y SIMETRICOS

14.48. Mostrar que para cualquier operador T, $T + T^*$ es autoadjunto y $T - T^*$ es anti-autoadjunto.

14.49. Supóngase que T es autoadjunto. Probar que $T^2(v) = 0$ implica $T(v) = 0$. Utilizar esto para demostrar que $T^n(v) = 0$ también implica $T(v) = 0$ para $n > 0$.

14.50. Sea V un espacio complejo con producto interno. Supóngase que $\langle T(v), v \rangle$ es real para todo $v \in V$. Probar que T es autoadjunto.

14.51. Supóngase que S y T son autoadjuntos. Mostrar que ST es autoadjunto si y sólo si S y T conmutan, es decir, $ST = TS$.

14.52. Para cada una de las siguientes matrices simétricas A, encontrar una matriz ortogonal P para la que $P^T AP$ sea diagonal:

$$a) \; A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}, \qquad b) \; A = \begin{pmatrix} 5 & 4 \\ 4 & -1 \end{pmatrix}, \qquad c) \; A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$$

14.53. Hallar una transformación ortogonal de coordenadas que diagonalice cada una de las siguientes formas cuadráticas: *a)* $q(x, y) = 2x^2 - 6xy + 10y^2$, *b)* $q(x, y) = x^2 + 8xy - 5y^2$.

14.54. Hallar una transformación ortogonal de coordenadas que diagonalice la forma cuadrática

$$q(x, y, z) = 2xy + 2xz + 2yz.$$

OPERADORES Y MATRICES NORMALES

14.55. Comprobar que $A = \begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}$ es normal. Encontrar una matriz unitaria P tal que P^*AP sea diagonal y hallar P^*AP.

14.56. Probar que una matriz triangular es normal si y sólo si es diagonal.

14.57. Demostrar que si T es normal en V, $\|T(v)\| = \|T^*(v)\|$ para todo $v \in V$. Demostrar que el recíproco se cumple en espacios complejos con producto interno.

14.58. Mostrar que los operadores autoadjuntos, anti-autoadjuntos y unitarios (ortogonales) son normales.

14.59. Supóngase que T es normal. Demostrar que:

a) T es autoadjunto si y sólo si sus valores propios son reales.

b) T es unitario si y sólo si sus valores propios tienen valor absoluto 1.

c) T es positivo si y sólo si sus valores propios son números reales no negativos.

14.60. Probar que si T es normal, T y T^* tienen el mismo núcleo y la misma imagen.

14.61. Supóngase que S y T son normales y conmutan. Mostrar que $S + T$ y ST también son normales.

14.62. Supóngase que T es normal y conmuta con S. Mostrar que T conmuta también con S^*.

PROBLEMAS SOBRE ISOMORFISMOS ENTRE ESPACIOS CON PRODUCTO INTERNO

14.63. Sea $S = \{u_1, \ldots, u_n\}$ una base ortonormal de un espacio con producto interno V sobre K. Probar que la aplicación $v \mapsto [v]_S$ es un isomorfismo (entre espacios con producto interno) entre V y K^n. (Aquí $[v]_S$ denota el vector coordenado de v en la base S.)

14.64. Mostrar que dos espacios con producto interno V y W sobre K son isomorfos si y sólo si V y W tienen la misma dimensión.

14.65. Supóngase que $\{u_1, \ldots, u_n\}$ y $\{u'_1, \ldots, u'_n\}$ son bases ortonormales de V y W, respectivamente. Sea $T: V \to W$ la aplicación lineal definida según $T(u_i) = u'_i$ para cada i. Mostrar que T es un isormorfismo.

14.66. Sea V un espacio con producto interno. Recuérdese (pág. 505) que cada $u \in V$ determina un funcional lineal $\hat{u}$ en el espacio dual V^* mediante la definición $\hat{u}(v) = \langle v, u \rangle$ para todo $v \in V$. Mostrar que la aplicación $u \mapsto \hat{u}$ es lineal y no singular y por ende un isomorfismo de V sobre V^*.

PROBLEMAS VARIOS

14.67. Probar que existe una base ortonormal $\{u_1, \ldots, u_n\}$ de V que consiste en vectores propios de T si y sólo si existen proyecciones ortogonales $E_1, \ldots, E_r$ y escalares $\lambda_1, \ldots, \lambda_r$ tales que:

$$\text{i) } T = \lambda_1 E_1 + \cdots + \lambda_r E_r\text{; ii) } E_1 + \cdots + E_r = I\text{; iii) } E_i E_j = 0 \text{ para } i \neq j$$

14.68. Supóngase que $V = U \oplus W$ y que $T_1 : U \to V$ y $T_2 : W \to V$ son lineales. Probar que $T = T_1 \oplus T_2$ es también lineal. Aquí T se define como sigue: Si $v \in V$ y $v = u + w$, donde $u \in U$, $w \in W$,

$$T(v) = T_1(u) + T_2(w)$$

14.69. Supóngase que U es un operador ortogonal en $\mathbf{R}^3$ con determinante positivo. Probar que U es bien una rotación o bien una reflexión en un plano.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

14.25. *a)* $\begin{pmatrix} 5+2i & 4+6i \\ 3-7i & 8-3i \end{pmatrix}$, *b)* $\begin{pmatrix} 3 & -i \\ -5i & 2i \end{pmatrix}$, *c)* $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$

14.26. $T^*(x, y, z) = (x + 3y, 2x + z, -4y)$.

14.27. $T^*(x, y, z) = (-ix + 3y, (2 - 3i)x + (2 + 5i)z, (3 + i)y - iz)$.

14.28. *a)* $u = (1, 2, -3)$, *b)* $u = (-i, 2 - 3i, 1 + 2i)$, *c)* $u = (18t^2 - 8t + 13)/15$.

14.32. *a)* $\begin{pmatrix} 2/\sqrt{13} & 3/\sqrt{13} \\ 3/\sqrt{13} & -2/\sqrt{13} \end{pmatrix}$, *b)* $\begin{pmatrix} 1/\sqrt{3} & (1-i)/\sqrt{3} \\ (1+i)/\sqrt{3} & -1/\sqrt{3} \end{pmatrix}$, *c)* $\begin{pmatrix} \frac{1}{2} & \frac{1}{2}i & \frac{1}{2} - \frac{1}{2}i \\ i/\sqrt{2} & -1/\sqrt{2} & 0 \\ \frac{1}{2} & -\frac{1}{2}i & -\frac{1}{2} + \frac{1}{2}i \end{pmatrix}$

14.45. Sólo i) y v) son positivas. Más aún, v) es definida positiva.

14.52. *a)* $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, *b)* $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, *c)* $P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}$

14.53. *a)* $x = (3x' - y')/\sqrt{10}$, $y = (x' + 3y')/\sqrt{10}$, *b)* $x = (2x' - y')/\sqrt{5}$, $y = (x' + 2y')/\sqrt{5}$.

14.54. $x = x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6}$, $y = x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6}$, $z = x'/\sqrt{3} - 2z'/\sqrt{6}$.

14.55. $P = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, $P^*AP = \begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix}$

APENDICE

Conjuntos y relaciones

CONJUNTOS, ELEMENTOS

Cualquier colección o lista de objetos bien definidos se llama un *conjunto*; los objetos que forman el conjunto se llaman sus *elementos* o *miembros*. Escribimos

$$p \in A \quad \text{si } p \text{ es un elemento del conjunto } A$$

Si todo elemento de A pertenece también a un conjunto B, es decir, si $x \in A$ implica $x \in B$, entonces llamamos a A un *subconjunto* de B o se dice que A está *contenido* en B; esto se denota por

$$A \subset B \quad \text{o} \quad B \supset A$$

Dos conjuntos son *iguales* si tienen los mismos elementos; esto es,

$$A = B \quad \text{si y sólo si} \quad A \subset B \quad \text{y} \quad B \subset A$$

Las negaciones de $p \in A$, $A \subset B$ y $A = B$ se escriben $p \notin A$, $A \not\subset B$ y $A \neq B$, respectivamente.

Especificamos un conjunto particular escribiendo todos sus elementos o estableciendo las propiedades que caracterizan los elementos de dicho conjunto. Por ejemplo,

$$A = \{1, 3, 5, 7, 9\}$$

significa que A es el conjunto formado por los números 1, 3, 5, 7 y 9; y

$$B = \{x : x \text{ es un número primo, } x < 15\}$$

significa que B es el conjunto de los números primos menores que 15. También empleamos símbolos especiales para representar los conjuntos que aparecen muy frecuentemente en el texto. A menos que específicamente se establezca lo contrario:

$\mathbf{N}$ = el conjunto de los enteros positivos: 1, 2, 3, ...;
$\mathbf{Z}$ = el conjunto de los enteros: ..., -2, -1, 0, 1, 2, ...;
$\mathbf{Q}$ = el conjunto de los números racionales;
$\mathbf{R}$ = el conjunto de los números reales;
$\mathbf{C}$ = el conjunto de los números complejos.

También usamos el símbolo $\emptyset$ para denotar el conjunto *vacío* o *nulo*, esto es, el conjunto que no contiene elementos; suponemos que este conjunto es un subconjunto de cualquier conjunto.

Frecuentemente los elementos de un conjunto son a su vez conjuntos. Por ejemplo, cada recta en un conjunto de rectas es un conjunto de puntos. Para ayudar a clarificar estas situaciones usamos las palabras *clase*, *colección* y *familia* como sinónimos con conjunto. Las palabras *subclase*, *subcolección* y *subfamilia* tienen significado análogo al de subconjunto.

EJEMPLO 1. Los conjuntos A y B anteriores también se pueden escribir así

$$A = \{x \in \mathbf{N} : x \text{ es impar}, \quad x < 10\} \quad \text{y} \quad B = \{2, 3, 5, 7, 11, 13\}$$

Obsérvese que $9 \in A$, pero $9 \notin B$, y $11 \in B$, pero $11 \notin A$; mientras que $3 \in A$ y $3 \in B$, y $6 \notin A$ y $6 \notin B$.

EJEMPLO 2. Los conjuntos de números se relacionan de la siguiente manera: $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$.

EJEMPLO 3. Sea $C = \{x : x^2 = 4, x \text{ es impar}\}$. Entonces $C = \emptyset$, esto es, C es el conjunto vacío.

EJEMPLO 4. Los elementos de la clase $\{\{2, 3\}, \{2\}, \{5, 6\}\}$ son los conjuntos $\{2, 3\}$, $\{2\}$ y $\{5, 6\}$.

Tenemos ahora el

Teorema 1: Sean A, B y C conjuntos arbitrarios. Entonces: i) $A \subset A$; ii) si $A \subset B$ y $B \subset A$, entonces $A = B$, iii) si $A \subset B$ y $B \subset C$, entonces $A \subset C$.

Recalcamos que $A \subset B$ no excluye la posibilidad de que $A = B$. Sin embargo, si $A \subset B$ y $A \neq B$, decimos entonces que A es un *subconjunto propio* de B. (Algunos autores usan el símbolo $\subseteq$ para un subconjunto y el símbolo $\subset$ únicamente para un subconjunto propio.)

Cuando hablamos de un *conjunto indizado* $\{a_i : i \in I\}$, o simplemente $\{a_i\}$, significamos que existe una aplicación ϕ del conjunto I en un conjunto A y que la imagen $\phi(i)$ de $i \in I$ se representa por a_i. El conjunto I se llama el *conjunto de índices* y se dice que los elementos a_i (el recorrido de ϕ) están indizados por I. Un conjunto $\{a_1, a_2, \ldots\}$ indizado por los enteros positivos $\mathbf{N}$ se llama una *sucesión*. Una clase indizada de conjuntos $\{A_i : i \in I\}$, o simplemente $\{A_i\}$, tiene un significado análogo, salvo que ahora la aplicación ϕ asigna a cada $i \in I$ un conjunto A_i en lugar de un elemento a_i.

OPERACIONES ENTRE CONJUNTOS

Sean A y B dos conjuntos arbitrarios. La *unión* de A y B, escrita $A \cup B$, es el conjunto de los elementos que pertenecen a A o a B; la *intersección* de A y B, escrita $A \cap B$, es el conjunto de los elementos que pertenecen a A y B simultáneamente:

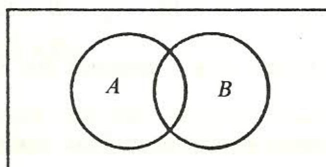
$$A \cup B = \{x : x \in A \quad \text{o} \quad x \in B\} \qquad \text{y} \qquad A \cap B = \{x : x \in A \quad \text{y} \quad x \in B\}$$

Si $A \cap B = \emptyset$, esto es, si A y B no tienen elementos en común, se dice entonces que A y B son *disjuntos*.

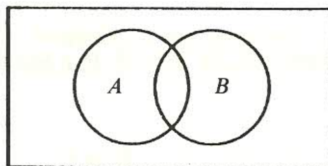
Suponemos que todos nuestros conjuntos son subconjuntos de un *conjunto universal* fijo (denotado aquí por U). El *complemento* de A, escrito A^c, es el conjunto de los elementos que no pertenecen a A:

$$A^c = \{x \in U : x \notin A\}$$

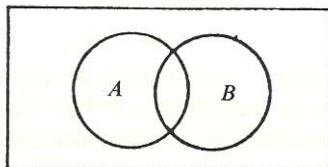
EJEMPLO 5. Los diagramas siguientes, llamados diagramas de Venn, ilustran las operaciones anteriores entre conjuntos. Los conjuntos se representan por áreas planas simples y U, el conjunto universal, por el área total del rectángulo.



$A \cup B$ está sombreado



$A \cap B$ está sombreado



A^c está sombreado

Los conjuntos con las operaciones anteriores satisfacen varias leyes o identidades que están escritas en la tabla siguiente. En efecto, tenemos el

Teorema 2: Los conjuntos satisfacen las leyes enunciadas en la Tabla 1.

Tabla 1

LEYES DEL ALGEBRA DE CONJUNTOS	
Leyes de idempotencia	
1a. $A \cup A = A$	1b. $A \cap A = A$
Leyes asociativas	
2a. $(A \cup B) \cup C = A \cup (B \cup C)$	2b. $(A \cap B) \cap C = A \cap (B \cap C)$
Leyes conmutativas	
3a. $A \cup B = B \cup A$	3b. $A \cap B = B \cap A$
Leyes distributivas	
4a. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	4b. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Leyes de identidad	
5a. $A \cup \emptyset = A$	5b. $A \cap U = A$
6a. $A \cup U = U$	6b. $A \cap \emptyset = \emptyset$
Leyes de complemento	
7a. $A \cup A^c = U$	7b. $A \cap A^c = \emptyset$
8a. $(A^c)^c = A$	8b. $U^c = \emptyset$, $\emptyset^c = U$
Leyes de De Morgan	
9a. $(A \cup B)^c = A^c \cap B^c$	9b. $(A \cap B)^c = A^c \cup B^c$

Observación: Cada una de las leyes anteriores se deduce de una ley lógica análoga. Por ejemplo,

$$A \cap B = \{x : x \in A \quad \text{y} \quad x \in B\} = \{x : x \in B \quad \text{y} \quad x \in A\} = B \cap A$$

(Aquí aprovechamos el hecho que la proposición compuesta «p y q», escrita $p \wedge q$, es lógicamente equivalente a la proposición compuesta «q y p», esto es, $q \wedge p$.)

La siguiente es la relación entre la inclusión de conjuntos y las operaciones anteriores.

Teorema 3: Cada una de las condiciones siguientes es equivalente a $A \subset B$:

i) $A \cap B = A$.
ii) $A \cup B = B$.
iii) $B^c \subset A^c$.
iv) $A \cap B^c = \emptyset$.
v) $B \cup A^c = U$.

Generalizamos las operaciones anteriores entre conjuntos de la siguiente manera. Sea $\{A_i : i \in I\}$ una familia arbitraria de conjuntos. La *unión* de los A_i, escrita $\bigcup_{i \in I} A_i$ (o simple-

mente $\bigcup_i A_i$), es el conjunto de los elementos que pertenecen por lo menos a uno de los A_i; y la *intersección* de los A_i, escrita $\bigcap_{i \in I} A_i$ (o simplemente $\bigcap_i A_i$), es el conjunto de los elementos que pertenecen a todos los A_i simultáneamente.

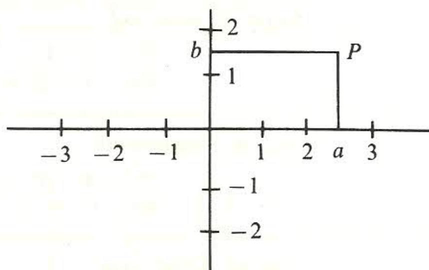
PRODUCTO CARTESIANO DE CONJUNTOS

Sean A y B dos conjuntos. El *producto cartesiano* de los conjuntos A y B, denotado por $A \times B$, está formado por todos los pares ordenados (a, b), donde $a \in A$ y $b \in B$:

$$A \times B = \{(a, b): a \in A, b \in B\}$$

El producto cartesiano de un conjunto por sí mismo, por ejemplo $A \times A$, se denota por A^2.

EJEMPLO 6. El lector está familiarizado con el plano cartesiano $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$, como se muestra a continuación. Cada punto P representa una pareja ordenada (a, b) de números reales y viceversa.



EJEMPLO 7. Sea $A = \{1, 2, 3\}$ y $B = \{a, b\}$. Entonces

$$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$

Nota: La pareja ordenada (a, b) se define rigurosamente por $(a, b) \equiv \{\{a\}, \{a, b\}\}$. De esta definición puede probarse la propiedad de «orden», esto es, $(a, b) = (c, d)$ si y sólo si $a = c$ y $b = d$.

El concepto de producto cartesiano de conjuntos puede extenderse a cualquier número finito de conjuntos en una forma natural. El conjunto *producto cartesiano* de los conjuntos $A_1, \ldots, A_m$, escrito $A_1 \times A_2 \times \cdots \times A_m$, es el conjunto formado por todas las m-uplas $(a_1, a_2, \ldots, a_m)$, donde $a_i \in A_i$ para cada i.

RELACIONES

Una *relación binaria*, o simplemente una *relación* R de un conjunto A a un conjunto B, asigna a cada pareja ordenada $(a, b) \in A \times B$ exactamente una de las siguientes proposiciones:

i) «a está relacionada con b», escrita $a R b$.
ii) «a no está relacionada con b», escrita $a \not{R} b$.

Una relación de un conjunto A al mismo conjunto A se llama una *relación* en A.

EJEMPLO 8. La inclusión de conjuntos es una relación en cualquier clase de conjuntos. Puesto que dado cualquier par de conjuntos A y B, o bien $A \subset B$ o $A \not\subset B$.

Observamos que cualquier relación R de A a B define un subconjunto único $\hat{R}$ de $A \times B$ de la siguiente manera:

$$\hat{R} = \{(a, b) : a R b\}$$

Recíprocamente, cualquier subconjunto $\hat{R}$ de $A \times B$ define una relación de A a B de la siguiente manera:

$$a R b \quad \text{si y sólo si} \quad (a, b) \in \hat{R}$$

En virtud de la correspondencia anterior entre las relaciones de A a B y los subconjuntos de $A \times B$, volvemos a definir una relación en la siguiente forma:

Definición: Una relación R de A a B es un subconjunto de $A \times B$.

RELACIONES DE EQUIVALENCIA

Una relación en un conjunto A se llama una *relación de equivalencia* si satisface los siguientes axiomas:

$|E_1|$ Cualquier $a \in A$ está relacionado consigo mismo.

$|E_2|$ Si a está relacionado con b, entonces b está relacionado con a.

$|E_3|$ Si a está relacionado con b y b está relacionado con c, entonces a está relacionado con c.

En general se dice que una relación es *reflexiva* si satisface $|E_1|$, *simétrica* si satisface $|E_2|$ y *transitiva* si satisface $|E_3|$. En otras palabras, una relación es una relación de equivalencias si es reflexiva, simétrica y transitiva.

EJEMPLO 9. Consideremos la relación $\subset$ de inclusión de conjuntos. Por el Teorema 1, $A \subset A$ para todo conjunto A; y si $A \subset B$ y $B \subset C$, entonces $A \subset C$, esto es, $\subset$ es reflexiva y transitiva. Por otra parte, $\subset$ no es simétrica, ya que $A \subset B$ y $A \neq B$ implica $B \not\subset A$.

EJEMPLO 10. En geometría euclidiana, la similaridad de triángulos es una relación de equivalencia. Pues si α, β y γ son triángulos arbitrarios, entonces: i) α es similar a sí mismo; ii) si α es similar a β, entonces β es similar a α; iii) si α es similar a β y β similar a γ, entonces α es similar a γ.

Si R es una relación de equivalencia en A, entonces la *clase de equivalencia* de cualquier elemento $a \in A$, denotada por $[a]$, es el conjunto de los elementos con los cuales a está relacionado:

$$[a] = \{x : a R x\}$$

La colección de las clases de equivalencia, denotada por A/R, se llama el *cociente* de A por R:

$$A/R = \{[a] : a \in A\}$$

La siguiente es la propiedad fundamental de las relaciones de equivalencia:

Teorema 4: Sea R una relación de equivalencia en A. Entonces el conjunto cociente A/R es una *partición* de A, esto es, cada $a \in A$ pertenece a un miembro de A/R y los miembros de A/R son disjuntos dos a dos.

EJEMPLO 11. Sea R_5 la relación en $\mathbf{Z}$, el conjunto de los enteros definido por

$$x \equiv y \pmod 5$$

que se lee «x es congruente a y módulo 5» y lo cual significa que «$x - y$ es divisible por 5». Entonces R_5 es una relación de equivalencia en $\mathbf{Z}$. Hay exactamente cinco clases diferentes de equivalencia en $\mathbf{Z}/R_5$:

$$A_0 = \{\ldots, -10, -5, 0, 5, 10\}$$
$$A_1 = \{\ldots, -9, -4, 1, 6, 11\}$$
$$A_2 = \{\ldots, -8, -3, 2, 7, 12\}$$
$$A_3 = \{\ldots, -7, -2, 3, 8, 13\}$$
$$A_4 = \{\ldots, -6, -1, 4, 9, 14\}$$

Como cada entero x se puede expresar en forma única, $x = 5q + r$, donde $0 \leqq r < 5$; observamos que $x \in A_r$, donde r es el residuo. Se puede ver que las clases de equivalencia son disjuntas dos a dos y que $\mathbf{Z} = A_0 \cup A_1 \cup A_2 \cup A_3 \cup A_4$.

Estructuras algebraicas

INTRODUCCION

Definimos aquí las estructuras algebraicas que aparecen en casi todas las ramas de la matemática. En particular, definimos lo que es un *cuerpo* (o *campo*), concepto que aparece en la definición de un espacio vectorial. Empezamos definiendo *grupo*, que es una estructura algebraica relativamente simple con únicamente una operación y que se usa como una base para muchos otros sistemas algebraicos.

GRUPOS

Sea G un conjunto no vacío con una operación binaria, esto es, a cada par de elementos $a, b \in G$ se le asigna un elemento $ab \in G$. Entonces G se llama un *grupo* si se cumplen los siguientes axiomas:

$|G_1|$ Para todo $a, b, c \in G$, tenemos $(ab)c = a(bc)$ (la *ley asociativa*).

$|G_2|$ Existe un elemento $e \in G$, llamado el elemento *identidad*, tal que $ae = ea = a$ para todo $a \in G$.

$|G_3|$ Para cada $a \in G$ existe un elemento $a^{-1} \in G$, llamado el *inverso* de a, tal que $aa^{-1} = a^{-1}a = e$.

Se dice que un grupo G es *abeliano* (o también *conmutativo*) si cumple la ley *conmutativa*, esto es, si $ab = ba$ para todo $a, b \in G$.

Cuando la operación binaria se representa por yuxtaposición como antes, se dice que el grupo G está escrito *multiplicativamente*. Algunas veces, cuando G es abeliano, la operación binaria se representa por $+$ y se dice que G está escrito *aditivamente*. En tal caso, el elemento identidad

se denota por 0 y se llama el elemento *cero*; y el inverso se denota por $-a$ y se llama el *negativo* de a.

Si A y B son subconjuntos de un grupo G, entonces escribimos

$$AB = \{ab : a \in A,\ b \in B\} \quad \text{o} \quad A + B = \{a + b : a \in A,\ b \in B\}$$

También escribimos a por $\{a\}$.

Un subconjunto H de un grupo G se llama un *subgrupo* de G si él mismo es un grupo para la operación de G. Si H es un subgrupo de G y $a \in G$, entonces el conjunto Ha se llama un *cogrupo a derecha* de H y el conjunto aH se llama un *cogrupo a izquierda* de H.

Definición: Un subgrupo H de G se llama un subgrupo *normal* si $a^{-1}Ha \subset H$ para todo $a \in G$. En forma equivalente, H es normal si $aH = Ha$ para todo $a \in G$, esto es, si los cogrupos a derecha y a izquierda de H coinciden.

Observamos que todo subgrupo de un grupo abeliano es normal.

Teorema 1: Sea H un subgrupo normal de G. Entonces los cogrupos de H en G forman un grupo para la multiplicación de cogrupos. Este grupo se llama el *grupo cociente* y se denota por G/H.

EJEMPLO 1. El conjunto $\mathbf{Z}$ de los enteros forma un grupo abeliano para la adición. (Observamos que los enteros pares forman un subgrupo de $\mathbf{Z}$ pero lo enteros impares no.) Sea H el conjunto de los múltiplos de 5, esto es, $H = \{\ldots, -10, -5, 0, 5, 10, \ldots\}$. Entonces H es un subgrupo (necesariamente normal) de $\mathbf{Z}$. Los cogrupos de H en $\mathbf{Z}$ son

$$\begin{aligned}
\bar{0} &= 0 + H = H = \{\ldots, -10, -5, 0, 5, 10, \ldots\} \\
\bar{1} &= 1 + H = \{\ldots, -9, -4, 1, 6, 11, \ldots\} \\
\bar{2} &= 2 + H = \{\ldots, -8, -3, 2, 7, 12, \ldots\} \\
\bar{3} &= 3 + H = \{\ldots, -7, -2, 3, 8, 13, \ldots\} \\
\bar{4} &= 4 + H = \{\ldots, -6, -1, 4, 9, 14, \ldots\}
\end{aligned}$$

Para cualquier otro entero $n \in \mathbf{Z}$, $\bar{n} = n + H$ coincide con uno de los cogrupos anteriores. Luego, por el teorema anterior, $\mathbf{Z}/H = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ forma un grupo para la adición de cogrupos; la siguiente es la tabla de adición:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

Este grupo cociente $\mathbf{Z}/H$ se conoce como el de los enteros módulo 5 y se representa frecuentemente por $\mathbf{Z}_5$. Análogamente, para todo entero positivo n existe el grupo cociente $\mathbf{Z}_n$, llamado los enteros módulo n.

EJEMPLO 2. Las permutaciones de n símbolos forman un grupo para la composición de aplicaciones; se llama el *grupo simétrico* de grado n y se denota por S_n. Investigamos S_3; sus elementos son

$$\varepsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad \phi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \qquad \phi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Aquí $\begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix}$ es la permutación que aplica $1 \mapsto i$, $2 \mapsto j$, $3 \mapsto k$. La tabla de multiplicación de S_3 es

	ε	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
ε	ε	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
σ_1	σ_1	ε	ϕ_1	ϕ_2	σ_2	σ_3
σ_2	σ_2	ϕ_2	ε	ϕ_1	σ_3	σ_1
σ_3	σ_3	ϕ_1	ϕ_2	ε	σ_1	σ_2
ϕ_1	ϕ_1	σ_3	σ_1	σ_2	ϕ_2	ε
ϕ_2	ϕ_2	σ_2	σ_3	σ_1	ε	ϕ_1

(El elemento en la fila a-ésima y columna b-ésima es ab.) El conjunto $H = \{\varepsilon, \sigma_1\}$ es un subgrupo de S_3; sus cogrupos a derecha y a izquierda son

Cogrupos a derecha	Cogrupos a izquierda
$H = \{\varepsilon, \sigma_1\}$	$H = \{\varepsilon, \sigma_1\}$
$H\phi_1 = \{\phi_1, \sigma_2\}$	$\phi_1 H = \{\phi_1, \sigma_3\}$
$H\phi_2 = \{\phi_2, \sigma_3\}$	$\phi_2 H = \{\phi_2, \sigma_2\}$

Observamos que los cogrupos a derecha y a izquierda son diferentes; por tanto, H no es un subgrupo normal de S_3.

Una aplicación f de un grupo G en un grupo G' se llama un *homomorfismo* si $f(ab) = f(a)f(b)$ para todo $a, b \in G$. (Si f es también biyectiva, esto es, uno-a-uno y sobre, entonces f se llama un *isomorfismo* y G y G' se dice que son *isomorfos*. Si $f : G \to G'$ es un homomorfismo, entonces el *núcleo* de f es el conjunto de elementos de G que se aplican en el elemento identidad $e' \in G'$:

$$\text{núcleo de } f = \{a \in G : f(a) = e'\}$$

(Como es corriente, $f(G)$ se llama la *imagen* de la aplicación $f : G \to G'$.) Tenemos ahora el

Teorema 2: Sea $f : G \to G'$ un homomorfismo con núcleo K. Entonces K es un subgrupo normal de G y el grupo cociente G/K es isomorfo a la imagen de f.

EJEMPLO 3. Sea G el grupo de los números reales para la adición y sea G' el grupo de números reales positivos para la multiplicación. La aplicación $f: G \to G'$ definida por $f(a) = 2^a$ es un homomorfismo ya que

$$f(a + b) = 2^{a+b} = 2^a 2^b = f(a)f(b)$$

En particular, f es biyectiva; luego G y G' son isomorfos.

EJEMPLO 4. Sea G el grupo de los números complejos diferentes de cero para la multiplicación y sea G' el grupo de los números reales diferentes de cero para la multiplicación. La aplicación $f: G \to G'$ definida por $f(z) = |z|$ es un homomorfismo ya que

$$f(z_1 z_2) = |z_1 z_2| = |z_1||z_2| = f(z_1)f(z_2)$$

El núcleo K de f está formado por los números complejos z sobre el círculo unitario, esto es, tales que $|z| = 1$. Luego G/K es isomorfo a la imagen de f, es decir, al grupo de los números reales positivos para la multiplicación.

ANILLOS, DOMINIOS DE INTEGRIDAD Y CUERPOS

Sea R un conjunto no vacío con dos operaciones binarias, una operación de adición (denotada por $+$) y una operación de multiplicación (denotada por yuxtaposición). Entones R se llama un *anillo* si cumple los siguientes axiomas:

[R_1] Para todo $a, b, c \in R$ tenemos $(a + b) + c = a + (b + c)$.

[R_2] Existe un elemento $0 \in R$, llamado el elemento *cero*, tal que $a + 0 = 0 + a = a$ para todo $a \in R$.

[R_3] Para cada $a \in R$ existe un elemento $-a \in R$, llamado el *negativo* de a, tal que $a + (-a) = (-a) + a = 0$.

[R_4] Para todo $a, b \in R$ tenemos $a + b = b + a$.

[R_5] Para todo $a, b, c \in R$ tenemos $(ab)c = a(bc)$.

[R_6] Para todo $a, b, c \in R$ tenemos
i) $a(b + c) = ab + ac$, ii) $(b + c)a = ba + ca$.

Observamos que los axiomas [R_1] hasta [R_4] pueden resumirse diciendo que R es un grupo abeliano para la adición.

La sustracción se define en R por $a - b \equiv a + (-b)$.

Se puede demostrar que (véase el Problema 25) $a \cdot 0 = 0 \cdot a = 0$ para todo $a \in R$.

R se llama un *anillo conmutativo* si $ab = ba$ para todo $a, b \in R$. Decimos también que R es un *anillo con elemento identidad* si existe un elemento diferente de cero $1 \in R$ tal que $a \cdot 1 = 1 \cdot a = a$ para todo $a \in R$.

Un subconjunto no vacío S de R se llama un *subanillo* de R si S mismo forma un anillo para las operaciones de R. Observamos que S es un subanillo de R si y sólo si $a, b \in S$ implica $a - b \in S$ y $ab \in S$.

Un subconjunto no vacío I de R se llama un *ideal a izquierda* en R si: i) $a - b \in I$ siempre que $a, b \in I$, ii) $ra \in I$ para $r \in R$, $a \in I$. Se observa que un ideal a izquierda I en R es también un subanillo de R. Similarmente podemos definir un *ideal a derecha* y un *ideal bilátero*.

Claramente, todos los ideales en un anillo conmutativo son biláteros. Empleamos la palabra ideal para hablar de los ideales biláteros a menos que específicamente se diga lo contrario.

Teorema 3: Sea I un ideal (bilátero) en un anillo R. Entonces los cogrupos $\{a + I : a \in R\}$ forman un anillo para la adición y multiplicación de cogrupos. Este anillo se denota por R/I y se llama el *anillo cociente*.

Ahora sea R un anillo conmutativo con elemento identidad. Para todo $a \in R$, el conjunto $(a) = \{ra : r \in R\}$ es un ideal; se llama el *ideal principal* generado por a. Si todo ideal en R es un ideal principal, entonces R se llama un *anillo de ideales principales*.

Definición: Un anillo conmutativo R con elemento identidad se llama un *dominio de integridad* si R no tiene *divisores de cero*, esto es, si $ab = 0$ implica $a = 0$ o $b = 0$.

Definición: Un anillo conmutativo R con elemento identidad se llama un *cuerpo* (o también un *campo*) si todo $a \in R$ diferente de cero tiene un *inverso multiplicativo*, esto es, existe un elemento $a^{-1} \in R$ tal que $aa^{-1} = a^{-1}a = 1$.

Un cuerpo es necesariamente un dominio de integridad, pues si $ab = 0$ y $a \neq 0$, entonces

$$b = 1 \cdot b = a^{-1}ab = a^{-1} \cdot 0 = 0$$

Hacemos notar que un cuerpo puede considerarse como un anillo conmutativo en el cual los elementos diferentes de cero forman un grupo para la multiplicación.

EJEMPLO 5. El conjunto $\mathbf{Z}$ de los enteros con las operaciones usuales de adición y multiplicación es el ejemplo clásico de un dominio de integridad con elemento identidad. Todo ideal I en $\mathbf{Z}$ es un ideal principal, esto es, $I = (n)$ para algún entero n. El anillo cociente $\mathbf{Z}_n = \mathbf{Z}/(n)$ se llama el *anillo de los enteros módulo n*. Si n es primo, entonces $\mathbf{Z}_n$ es un cuerpo. De otra parte, si n no es primo, entonces $\mathbf{Z}_n$ tiene divisores de cero. Por ejemplo, en el anillo $\mathbf{Z}_6$, $\bar{2}\,\bar{3} = \bar{0}$ y $\bar{2} \neq \bar{0}$ y $\bar{3} \neq \bar{0}$.

EJEMPLO 6. Los números racionales $\mathbf{Q}$ y los números reales $\mathbf{R}$, cada uno de ellos, forman un cuerpo para las operaciones corrientes de adición y multiplicación.

EJEMPLO 7. Sea $\mathbf{C}$ el conjunto de las parejas ordenadas de números reales con la adición y la multiplicación definidas por

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

Entonces $\mathbf{C}$ satisface todas las propiedades requeridas de un cuerpo. En efecto, $\mathbf{C}$ es precisamente el cuerpo de los números complejos.

EJEMPLO 8. El conjunto M de todas las matrices 2×2 con componentes reales forma un anillo no conmutativo con divisores de cero para las operaciones de adición y multiplicación de matrices.

EJEMPLO 9. Sea R un anillo cualquiera. Entonces el conjunto $R[x]$ de todos los polinomios sobre R forma un anillo con respecto a las operaciones corrientes de adición y multiplicación de polinomios. Además, si R es un dominio de integridad, entonces $R[x]$ también es un dominio de integridad.

Sea D un dominio de integridad. Decimos que b *divide* a a en D si $a = bc$ para algún $c \in D$. Un elemento $u \in D$ se llama una *unidad* si u divide a 1, esto es, si u tiene un inverso multiplicativo. Un elemento $b \in D$ se llama un *asociado* de $a \in D$ si $b = ua$ para alguna unidad $u \in D$. Un elemento no unidad $p \in D$ se dice *irreducible* si $p = ab$ implica que a o b son unidades.

Un dominio de integridad D se llama un *dominio de factorización única* si cualquier elemento no unidad $a \in D$ puede escribirse en forma única (salvo asociados y el orden de los elementos) como un producto de elementos irreducibles.

EJEMPLO 10. El anillo $\mathbf{Z}$ de los enteros es el ejemplo clásico de un dominio de factorización única. Las unidades de $\mathbf{Z}$ son 1 y -1. Los únicos asociados de $n \in \mathbf{Z}$ son n y $-n$. Los elementos irreducibles de $\mathbf{Z}$ son los números primos.

EJEMPLO 11. El conjunto $D = \{a + b\sqrt{13} : a, b \text{ enteros}\}$ es un dominio de integridad. Las unidades de D son ± 1, $18 \pm 5\sqrt{13}$ y $-18 \pm 5\sqrt{13}$. Los elementos 2, $3 - \sqrt{13}$ y $-3 - \sqrt{13}$ son irreducibles en D. Observemos que $4 = 2 \cdot 2 = (3 - \sqrt{13})(-3 - \sqrt{13})$. Por tanto, D no es un dominio de factorización única. (Véase el Problema 40.)

MODULOS

Sea M un conjunto no vacío y sea R un anillo con elemento identidad. Se dice que M es un R-*módulo* (a izquierda) si M es un grupo aditivo abeliano y existe una aplicación $R \times M \to M$ que satisface los siguientes axiomas:

$[M_1]$ $r(m_1 + m_2) = rm_1 + rm_2$

$[M_2]$ $(r + s)m = rm + sm$

$[M_3]$ $(rs)m = r(sm)$

$[M_4]$ $\mathbf{1} \cdot m = m$

para todo $r, s \in R$ y todo $m_i \in M$.

Recalcamos que un R-módulo es una generalización de un espacio vectorial donde únicamente pedimos a los escalares que formen un anillo en lugar de un cuerpo.

EJEMPLO 12. Sea G un grupo aditivo abeliano arbitrario. Hacemos de G un módulo sobre el anillo $\mathbf{Z}$ de los enteros definiendo

$$ng = \overbrace{g + g + \cdots + g}^{n \text{ veces}}, \qquad 0g = 0, \qquad (-n)g = -ng$$

donde n es cualquier entero positivo.

EJEMPLO 13. Sea R un anillo y sea I un ideal en R. Entonces I puede considerarse como un módulo sobre R.

EJEMPLO 14. Sea V un espacio vectorial sobre un cuerpo K y sea $T: V \to V$ una aplicación lineal. Hacemos de V un módulo sobre el anillo $K|x|$ de los polinomios sobre K definiendo $f(x)v = f(T)(v)$. El lector comprobará que la multiplicación por escalar queda bien definida.

Sea M un módulo sobre R. Un subgrupo aditivo N de M se llama un *submódulo* de M si $u \in N$ y $k \in R$ implican $ku \in N$. (Notamos que N es entonces un módulo sobre R.)

Sean M y M' dos R-módulos. Una aplicación $T: M \to M'$ se llama un *homomorfismo* (o también *R-homomorfismo* o *R-lineal*) si

$$\text{i) } T(u + v) = T(u) + T(v) \quad \text{y} \quad \text{ii) } T(ku) = kT(u)$$

para todo u, $v \in M$ y todo $k \in R$.

PROBLEMAS SUPLEMENTARIOS

GRUPOS

1. Determinar cuáles de los siguientes sistemas forman un grupo G:

i) G = conjunto de los enteros, operación sustracción.
ii) $G = \{1, -1\}$, operación multiplicación.
iii) G = conjunto de los números racionales diferentes de cero, operación división.
iv) G = conjunto de las matrices no singulares $n \times n$, operación multiplicación de matrices.
v) $G = \{a + bi : a, b \in \mathbf{Z}\}$, operación adición.

2. Mostrar que en un grupo G

i) el elemento identidad de G es único;
ii) cada $a \in G$ tiene un inverso único $a^{-1} \in G$;
iii) $(a^{-1})^{-1} = a$ y $(ab)^{-1} = b^{-1}a^{-1}$;
iv) $ab = ac$ implica $b = c$ y $ba = ca$ implica $b = c$.

3. En un grupo G, las potencias de $a \in G$ se definen por

$$a^0 = e,\ a^n = aa^{n-1},\ a^{-n} = (a^n)^{-1}, \text{ donde } n \in \mathbf{N}$$

Mostrar que las fórmulas siguientes se cumplen para enteros arbitrarios r, s, $t \in \mathbf{Z}$: i) $a^r a^s = a^{r+s}$, ii) $(a^r)^s = a^{rs}$, iii) $(a^{r+s})^t = a^{rt+st}$.

4. Mostrar que si G es un grupo abeliano, entonces $(ab)^n = a^n b^n$ para todo a, $b \in G$ y cualquier entero $n \in \mathbf{Z}$.

5. Supongamos que G es un grupo tal que $(ab)^2 = a^2 b^2$ para todo a, $b \in G$. Mostrar que G es abeliano.

6. Supongamos que H es un subconjunto de un grupo G. Mostrar que H es un subgrupo de G si y sólo si: i) H es no vacío, ii) $a, b \in H$ implica $ab^{-1} \in H$.

7. Demostrar que la intersección de cualquier número de subgrupo de G también es un subgrupo de G.

8. Mostrar que el conjunto de todas las potencias de $a \in G$ es un subgrupo de G; éste se llama el *grupo cíclico* generado por a.

9. Se dice que un grupo G es *cíclico* si G es generado por algún $a \in G$, esto es, $G = \{a^n : n \in \mathbf{Z}\}$. Mostrar que cualquier subgrupo de un grupo cíclico es cíclico.

10. Supongamos ue G es un grupo cíclico. Mostrar que G es isomorfo al conjunto $\mathbf{Z}$ de los enteros para la adición o al conjunto $\mathbf{Z}_n$ (de los enteros módulo n) para la adición.

11. Sea H un subgrupo de G. Mostrar que los cogrupos a derecha (a izquierda) de H particionan a G en subconjuntos disjuntos dos a dos.

12. El *orden* de un grupo G denotado por $|G|$ es el número de elementos de G. Demostrar el teorema de Lagrange: Si H es un subgrupo de un grupo finito G, entonces $|H|$ divide a $|G|$.

13. Supongamos que $|G| = p$, donde p es primo. Mostrar que G es cíclico.

14. Supongamos que H y N son dos subgrupos de G con N normal. Mostrar que: i) HN es un subgrupo de G, ii) $H \cap N$ es un subgrupo normal de G.

15. Sea H un subgrupo de G con dos cogrupos únicamente a derecha (a izquierda). Mostrar que H es un subgrupo normal de G.

16. Demostrar el Teorema 1: Sea H un subgrupo normal de G. Entonces los cogrupos de H en G forman un grupo G/H para la multiplicación de cogrupos.

17. Supongamos que G es un grupo abeliano. Mostrar que cualquier grupo factor G/H también es abeliano.

18. Sea $f: G \to G'$ un homomorfismo de grupos. Mostrar que

i) $f(e) = e'$, donde e y e' son los elementos idénticos de G y G' respectivamente;

ii) $f(a^{-1}) = f(a)^{-1}$ para todo $a \in G$.

19. Demostrar el Teorema 2: Sea $f: G \to G'$ un homomorfismo de grupos con núcleo K. Entonces K es un subgrupo normal de G y el grupo cociente G/K es isomorfo a la imagen de f.

20. Sea G el grupo multiplicativo de los números complejos z tales que $|z| = 1$ y sea $\mathbf{R}$ el grupo aditivo de los números reales. Probar que G es isomorfo a $\mathbf{R}/\mathbf{Z}$.

21. Para un elemento fijo $g \in G$ sea $\hat{g}: G \to G$ definida por $\hat{g}(a) = g^{-1}ag$. Mostrar que $\hat{g}$ es un isomorfismo de G sobre G.

22. Sea G el grupo multiplicativo de las matrices $n \times n$ no singulares sobre $\mathbf{R}$. Mostrar que la aplicación $A \mapsto |A|$ es un homomorfismo de G en el grupo multiplicativo de los números reales diferentes de cero.

23. Sea G un grupo abeliano. Para un elemento fijo $n \in \mathbf{Z}$, mostrar que la aplicación $a \mapsto a^n$ es un homomorfismo de G en G.

24. Supongamos que H y N son dos subgrupos de G con N normal. Probar que $H \cap N$ es normal en H y $H/(H \cap N)$ es isomorfo a HN/N.

ANILLOS

25. Mostrar que en un anillo R: i) $a \cdot 0 = 0 \cdot a = 0$, ii) $a(-b) = (-a)b = -ab$, iii) $(-a)(-b) = ab$.

26. Mostrar que en un anillo R con un elemento identidad: i) $(-1)a = -a$, ii) $(-1)(-1) = 1$.

27. Supongamos que $a^2 = a$ para todo $a \in R$. Probar que R es un anillo conmutativo. (Un anillo con esta propiedad se llama un *anillo booleano*.)

28. Sea R un anillo con elemento identidad. Formamos con R otro anillo $\hat{R}$ definiendo $a \oplus b = a + b + 1$ y $a \cdot b = ab + a + b$. i) Verificar que $\hat{R}$ es un anillo. ii) Determinar los elementos 0 y 1 de $\hat{R}$.

29. Sea G un grupo (aditivo) abeliano cualquiera. Definimos una multiplicación en G por $a \cdot b = 0$. Mostrar que esta multiplicación hace de G un anillo.

30. Probar el Teorema 3: Sea I un ideal (bilátero) en un anillo R. Entonces los cogrupos $\{a + I : a \in R\}$ forman un anillo para la adición y la multiplicación de cogrupos.

31. Sean I_1 e I_2 dos ideales en R. Probar que $I_1 + I_2$ e $I_1 \cap I_2$ son también ideales en R.

32. Sean R y R' dos anillos. Una aplicación $f : R \to R'$ se llama un *homomorfismo* (o también un *homomorfismo de anillos*) si

$$\text{i)} \quad f(a + b) = f(a) + f(b) \qquad \text{y} \qquad \text{ii)} \quad f(ab) = f(a)f(b)$$

para todo $a, b \in R$. Probar que si $f : R \to R'$ es un homomorfismo, entonces el conjunto $K = \{r \in R : f(r) = 0\}$ es un ideal en R. (El conjunto K se llama el *núcleo* de f.)

DOMINIOS DE INTEGRIDAD Y CUERPOS

33. Probar que en un dominio de integridad D si $ab = ac$, $a \neq 0$, entonces $b = c$.

34. Probar que $F = \{a + b\sqrt{2} : a, b \text{ racionales}\}$ es un cuerpo.

35. Probar que $D = \{a + b\sqrt{2} : a, b \text{ enteros}\}$ es un dominio de integridad pero no un cuerpo.

36. Probar que un dominio de integridad finito D es un cuerpo.

37. Mostrar que los únicos ideales en un cuerpo K son $\{0\}$ y K.

38. Un número complejo $a + bi$, donde a, b son enteros, se llama un *entero gaussiano*. Mostrar que el conjunto G de los enteros gaussianos es un dominio de integridad. También mostrar que las unidades en G son ± 1 y $\pm i$.

39. Sea D un dominio de integridad y sea I un ideal en D. Probar que el anillo factor D/I es un dominio de integridad si y sólo si I es un ideal primo. (Un ideal I es *primo* si $ab \in I$ implica $a \in I$ o $b \in I$.)

40. Consideremos el dominio integridad $D = \{a + b\sqrt{13} : a, b \text{ enteros}\}$ (véase el Ejemplo 11). Si $\alpha = a + b\sqrt{13}$, definimos $N(\alpha) = a^2 - 13b^2$. Probar: i) $N(\alpha\beta) = N(\alpha)N(\beta)$; ii) α es una unidad si y sólo si $N(\alpha) = \pm 1$; iii) las unidades de D son ± 1, $18 \pm 5\sqrt{13}$ y $-18 \pm 5\sqrt{13}$; iv) los números 2, $3 - \sqrt{13}$ y $-3 - \sqrt{13}$ son irreducibles.

MODULOS

41. Sea M un R-módulo y sean A y B dos submódulos de M. Mostrar que $A + B$ y $A \cap B$ son también submódulos de M.

42. Sea M un R-módulo con un submódulo N. Mostrar que los cogrupos $\{u + N : u \in M\}$ forman un R-módulo para la adición de cogrupos y la multiplicación por escalar definida por $r(u + N) = ru + N$. (Este módulo se denota por M/N y se llama el *módulo cociente*.)

43. Sean M y M' dos R-módulos y sea $f: M \to M'$ un R-homomorfismo. Mostrar que el conjunto $K = \{u \in M : f(u) = 0\}$ es un submódulo de f. (El conjunto K se llama el *núcleo* de f.)

44. Sea M un R-módulo y sea $E(M)$ el conjunto de todos los R-homomorfismo de M en sí mismo. Definir las operaciones apropiadas de adición y multiplicación en $E(M)$ para que sea un anillo.

Polinomios sobre un cuerpo

INTRODUCCION

El anillo $K[t]$ de los polinomios sobre un cuerpo K tiene muchas propiedades análogas a propiedades de los enteros. Estas juegan un papel importante en la obtención de formas canónicas para un operador lineal T en un espacio vectorial V sobre K.

Cualquier polinomio en $K[t]$ puede escribirse de la forma

$$f(t) = a_n t^n + \cdots + a_1 t + a_0$$

La entrada a_k se denomina el k-ésimo *coeficiente* de f. Si n es el mayor entero para el que $a_n \neq 0$, decimos que el grado de f es n, escrito

$$\text{grado } f = n$$

Asimismo, llamamos a a_n el *coeficiente dominante* de f, y si $a_n = 1$, decimos que f es un *polinomio normalizado*. Por otra parte, si todo coeficiente de f es 0, f se llama el *polinomio cero*, escrito $f = 0$. El grado del polinomio cero no está definido.

DIVISIBILIDAD. MAXIMO COMUN DIVISOR

El siguiente teorema formaliza el proceso conocido como «división larga».

Teorema 1 (Algoritmo de Euclides): Sean f y q polinomios sobre un cuerpo K con $g \neq 0$. Existen polinomios q y r tales que

$$f = qg + r$$

donde $r = 0$ o grado $r <$ grado g.

Demostración: Si $f = 0$ o si grado $f <$ grado g, tenemos la representación requerida

$$f = 0g + f$$

Supongamos ahora grado $f \geqslant$ grado g, digamos

$$f = a_n t^n + \cdots + a_1 t + a_0 \quad \text{y} \quad g = b_m t^m + \cdots + b_1 t + b_0$$

donde a_n, $b_m \neq 0$ y $n \geqslant m$. Construimos el polinomio

$$f_1 = f - \frac{a_n}{b_m} t^{n-m} g \tag{1}$$

Entonces grado $f_1 <$ grado f. Por inducción existen polinomios q_1 y r tales que

$$f_1 = q_1 g + r$$

donde $r = 0$ o grado $r <$ grado g. Sustituyendo esto en [1] y despejando f,

$$f = \left(q_1 + \frac{a_n}{b_m} t^{n-m}\right) g + r$$

que es la representación deseada.

Teorema 2: El anillo $K[t]$ de los polinomios sobre un cuerpo K es un anillo de ideales principales. Si I es un ideal en $K[t]$, existe un único polinomio normalizado d que genera I, esto es, tal que d es divisor de todo polinomio $f \in I$.

Demostración: Sea d un polinomio de grado mínimo en I. Dado que podemos multiplicar d por un escalar no nulo y permanecer todavía en I, podemos suponer sin pérdida de generalidad que d es un polinomio normalizado. Supongamos ahora $f \in I$. En virtud del Teorema 1, existen polinomios q y r tales que

$$f = qd + r, \qquad \text{donde } r = 0 \text{ o grado } r < \text{grado } d$$

Ahora f, $d \in I$ implica $qd \in I$ y por tanto $r = f - qd \in I$. Pero d es un polinomio de grado mínimo en I. En consecuencia, $r = 0$ y $f = qd$, es decir, d es divisor de f. Resta probar que d es único. Si d' es otro polinomio normalizado que genera I, entonces d es divisor de d' y d' es divisor de d. Esto implica $d = d'$, porque d y d' están normalizados. Queda así demostrado el teorema.

Teorema 3: Sean f y g polinomios no nulos en $K[t]$. Existe un único polinomio normalizado d tal que: i) d es divisor de f y g, ii) si d' es divisor de f y g, necesariamente lo es de d.

Definición: El polinomio d precedente se denomina el máximo común divisor de f y g. Si $d = 1$, se dice que f y g son primos entre sí.

Demostración del Teorema 3: El conjunto $I = \{mf + ng : m, n \in K[t]\}$ es un ideal. Sea d el polinomio normalizado que genera I. Nótese que $f, g \in I$, luego d es divisor de f y g. Supongamos ahora que d' es divisor de f y g. Sea J el ideal generado por d'. En ese caso, $f, g \in J$ y por consiguiente $I \subset J$. De acuerdo con esto, $d \in J$, de modo que d' es divisor de d como se pretendía. Falta probar que d es único. Si d_1 es otro máximo común divisor de f y g (normalizado), d es divisor de d_1 y d_1 lo es de d. Esto implica $d = d_1$, ya que d y d_1 están normalizados. Queda así demostrado el teorema.

Corolario 4: Sea d el máximo común divisor de los polinomios f y g. Existen polinomios m y n tales que $d = mf + ng$. En particular, si f y g son primos entre sí, existen polinomios m y n tales que $mf + ng = 1$.

El corolario deriva directamente del hecho de que d genera el ideal

$$I = \{mf + ng : m, n \in K[t]\}$$

FACTORIZACION

Se dice que un polinomio $p \in K[t]$ de grado positivo es irreducible si $p = fg$ implica que f o g es un escalar.

Lema 5: Supongamos que $p \in K[t]$ es irreducible. Si p es divisor del producto fg de polinomios $f, g \in K[t]$, p es divisor de f o de g. Con mayor generalidad, si p es divisor del producto de n polinomios $f_1 f_2 \dots f_n$, entonces es divisor de uno de ellos.

Demostración: Supongamos que p es divisor de fg pero no de f. Como p es irreducible, los polinomios f y p deben ser primos entre sí. Existen, pues, polinomios $m, n \in K[t]$ tales que $mf + np = 1$. Multiplicando esta ecuación por g obtenemos $mfg + npg = g$. Pero p es divisor de fg y por tanto de mfg y de npg; por esta razón, p es divisor de la suma $g = mfg + npg$.

Supongamos ahora que p es divisor de $f_1 f_2 \dots f_n$. Si p es divisor de f_1, el lema está demostrado. Si no, por el resultado anterior, p es divisor del producto $f_2 \dots f_n$. Por inducción en n, p es divisor de uno de los polinomios $f_2, \dots, f_n$. Queda así demostrado el lema.

Teorema 6 (Teorema de factorización única): Sea f un polinomio no nulo en $K[t]$. Entonces f puede escribirse de forma única (salvo orden) como un producto

$$f = kp_1 p_2 \dots p_n$$

donde $k \in K$ y los p_i son polinomios irreducibles normalizados en $K[t]$.

Demostración: Probamos primero la existencia de tal producto. Si f es irreducible o si $f \in K$, claramente existe un producto semejante. Por otra parte, supongamos $f = gh$, donde g y h no son escalares. En tal caso, g y h tienen grados menores que el de f. Por inducción podemos suponer

$$g = k_1 g_1 g_2 \dots g_r \quad \text{y} \quad h = k_2 h_1 h_2 \dots h_s$$

donde k_1, $k_2 \in K$ y los g_i y h_j son polinomios irreducibles normalizados. En consecuencia,

$$f = (k_1 k_2) g_1 g_2 \dots g_r h_1 h_2 \dots h_s$$

es nuestra representación deseada.

A continuación demostramos la unicidad (salvo orden) de tal producto para f. Supongamos

$$f = k p_1 p_2 \dots p_n = k' q_1 q_2 \dots q_m$$

donde k, $k' \in K$ y los $p_1, \dots, p_n, q_1, \dots, q_m$ son polinomios irreducibles normalizados. Ahora p_1 es divisor de $k' q_1 \dots q_m$. Como p_1 es irreducible, debe ser divisor de uno de los q_i, según el Lema 5. Sea p_1 divisor de, por ejemplo, q_1. Dado que p_1 y q_1 son irreducibles y están normalizados, $p_1 = q_1$. De acuerdo con ello,

$$k p_2 \dots p_n = k' q_2 \dots q_m$$

Por inducción tenemos que $n = m$ y $p_2 = q_2, \dots, p_n = q_m$ para alguna reordenación de los q_i. También tenemos que $k = k'$. Queda, pues, demostrado el teorema.

Si el cuerpo K es el cuerpo complejo $\mathbf{C}$, disponemos del siguiente

Teorema 7 (Teorema fundamental del álgebra): Sea $f(t)$ un polinomio no nulo sobre el cuerpo complejo $\mathbf{C}$. Entonces $f(t)$ puede escribirse de forma única (salvo orden) como un producto

$$f(t) = k(t - r_1)(t - r_2) \dots (t - r_n)$$

donde k, $r_i \in \mathbf{C}$, es decir, como un producto de polinomios lineales.

En el caso del cuerpo real $\mathbf{R}$, disponemos del siguiente resultado.

Teorema 8: Sea $f(t)$ un polinomio no nulo sobre el cuerpo real $\mathbf{R}$. Entonces $f(t)$ puede escribirse de forma única (salvo orden) como un producto

$$f(t) = k p_1(t) p_2(t) \dots p_m(t)$$

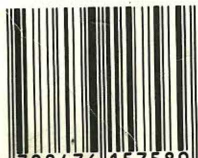
donde $k \in \mathbf{R}$ y los $p_i(t)$ son polinomios irreducibles normalizados de grado uno o dos.

Indice



OTRAS OBRAS DE INTERES PUBLICADAS POR McGRAW-HILL/INTERAMERICANA

ABELLANAS/GALINDO. *Métodos de Cálculo* (Schaum).
AMILLO/ARRIAGA. *Análisis matemático con aplicaciones a la computación.*
AYRES. *Algebra moderna* (Schaum).
AYRES. *Cálculo diferencial e integral* (3.ª ed.) (Schaum).
AYRES. *Ecuaciones diferenciales* (Schaum).
CHURCHILL. *Variable compleja y aplicaciones* (5.ª ed.).
GRAFE. *Matemáticas para economistas* (2.ª ed.).
GRANERO. *Algebra y geometría analítica.*
GRANERO. *Cálculo.*
KITCHEN. *Cálculo.*
LARSON. *Cálculo y geometría analítica* (3.ª ed.).
LIPSCHUTZ. *Probabilidad.*
MARCELLAN/CASASUS/ZARZO. *Ecuaciones diferenciales.*
RUDIN. *Análisis real y complejo.*
RUDIN. *Principios de análisis matemático* (3.ª ed.).
SIMMONS. *Ecuaciones diferenciales.*
SPIEGEL. *Estadística* (2.ª ed.) (Schaum).
SPIEGEL/ABELLANAS. *Fórmulas y tablas de matemática aplicada* (Schaum).
STEIN. *Cálculo y geometría analítica* (3.ª ed.).
TORREGROSA/JORDAN. *Algebra lineal y sus aplicaciones.*
WYLIE. *Matemáticas superiores para ingenieros* (2.ª ed.).





ISBN: 84-7615-758-4